

AD-A080 172

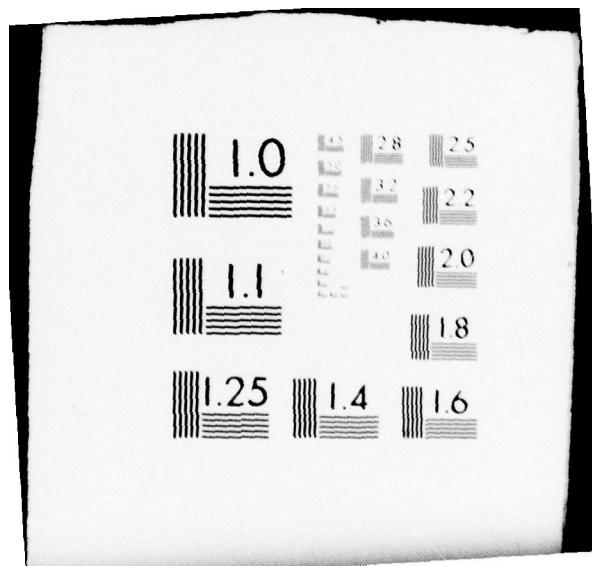
AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHOOL--ETC F/G 9/4
PERFORMANCE OF A SOURCE/CHANNEL ENCODED IMAGERY TRANSMISSION SY--ETC(U)
DEC 79 R A DURYEA
UNCLASSIFIED AFIT/GE/EE/79D-12

NL

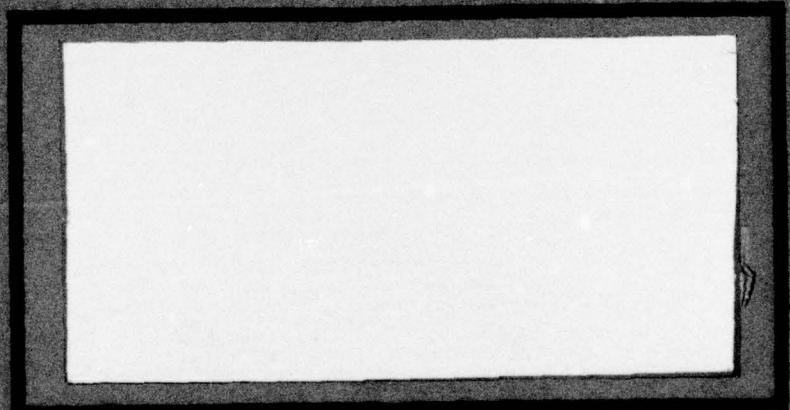
1 OF 2

AD
A080172





ADA080172



FILE COPY

UNITED STATES AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

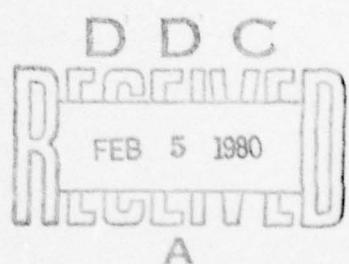
AFIT/GE/EE/79D-12

PERFORMANCE OF A SOURCE/CHANNEL ENCODED
IMAGERY TRANSMISSION SYSTEM

THESIS

AFIT/GE/EE/79D-12

Robert A. Duryea
Capt USAF



Approved for public release; distribution unlimited

14

AFIT/GE/EE/79D-12

⑥

PERFORMANCE OF A SOURCE/CHANNEL ENCODED
IMAGERY TRANSMISSION SYSTEM.

⑨

Master's THESIS

Accession For	
NTIS GRANT	
DDC TAB	
Unannounced	
Justification	
By	
Distribution/	
Availability Codes	
Disk	Available and/or special
A	

Presented to the faculty of the School of Engineering
of the Air Force Institute of Technology
Air University
in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

by

⑩

Robert A. Duryea, B.S., M.B.A.
Capt USAF
Graduate Electrical Engineering

⑪

December 1979

⑫ 159

Approved for public release; distribution unlimited

012 225

1B

Preface

In the last two decades there has been a great deal of research focused on finding efficient source encoding procedures for imagery. During that same period there has been a second large research effort to develop error correcting channel codes. This work attempts to show that the results of these previous research efforts can be combined into a source/channel encoded imagery transmission system which can efficiently transmit good quality imagery over even very noisy communication links.

My interest in this research was sparked 3 years ago while I was assigned to Rome Air Development Center (RADC). At the time I was a Project Engineer for an experimental image transmission system. The system was designed to transmit imagery over 3KHz bandwidth voice channels. One of the tests we attempted with the system was a transmission over the RADC TRC-97 Troposcatter Test Range. After 2 weeks of "tweaking" this permanently installed communication link, the best error rate achieved was 5×10^{-4} . At that rate we were able to transmit 1 usable image. Not until I came to AFIT and was enlightened by (then Capt and AFIT instructor) Gregg Vaughn about error correcting codes did I see any possibility of ever reliably transmitting imagery over such noisy communication links.

I would like to take this opportunity to recognize several people for their support during the course of this research. First, I would like to thank my advisor Major Joseph W. Carl for his guidance and for his patience in waiting for some concrete results to emerge from this research. I would also like to thank Debbie Galle who spent many hours typing this thesis. Finally I would like to thank my wife Marie and my son Andrew for their understanding during this time.

Robert A. Duryea

Contents

	<u>Page</u>
Preface	ii
List of Figures	v
List of Tables.	viii
Abstract.	ix
I. Introduction	1
Background	2
Organization of the Thesis	4
II. Rate Distortion Curve for an Imagery Source.	6
Image Model.	6
Distortion Measure	12
Image Rate Distortion Function	19
III. Performance of a Source Encoded Imagery Transmission System	31
Discrete Cosine Transform (DCT).	32
Block Quantization of the DCT Coefficients	42
Theoretical Performance of the Source Encoded Imagery Transmission System.	58
Calculated Performance of the Source Encoded Imagery Transmission System.	73
IV. Performance Improvement with Combined Source/Channel Encoding Procedures.	78
Channel Coding Procedures.	78
Application of Channel Encoding to the Block Quantizer Output	85
Calculated Performance of the Source/Channel Encoded Imagery Transmission System.	87
Summary of Source/Channel Coder Performance.	104
V. Image Transmission Simulation.	108
Simulation Hardware.	108
Simulation Software.	110
Simulation Results	116

Contents (continued)

	<u>Page</u>
VI. Conclusions and Recommendations	137
Performance Improvement Through the Application of Channel Codes.	137
Weighted Mean Square Error Versus Mean Square Error	138
Recommendations for Further Study	139
Bibliography	142
Vita	146

List of Figures

<u>Figure</u>		<u>Page</u>
2.1	Truck: Original Image and Histogram	8
2.2	SMSITE: Original Image and Histogram.	9
2.3	Bomber: Original Image and Histogram.	10
2.4	Power Spectral Density of Gauss-Markov Image Model .	13
2.5	Measured Contrast Threshold of the Human Visual System	16
2.6	Models for Human Visual System Transfer Function . .	18
2.7	Calculation Procedure for Mean Square Error Rate Distortion Function.	21
2.8	Calculation Procedure for Weighted Mean Square Error Rate Distortion Function	24
2.9	Rate Distortion Function for the 2-D Gauss-Markov Image Model.	26
2.10	Comparison of WMSE and MSE Rate Distortion Functions.	27
2.11	Dependence of R(D) on Average Lumination (B) for Amplitude Normalized WMSE.	29
2.12	Dependence of R(D) on Average Lumination (B) for Power Normalized WMSE.	30
3.1	Source/Channel Encoded Digital Image Transmission System	33
3.2	Performance of Various Transforms in Encoding a 1-D Gauss-Markov Source.	35
3.3	Fast Discrete Cosine Transform Signal Flow Graph . .	38
3.4	Weighted Mean Square Error in the Spatial Domain . .	46
3.5	Weighted Mean Square Error in the Cosine Transform Domain	47
3.6	Comparison of Rate Distortion Functions for Image and DCT Sources.	52

List of Figures (continued)

<u>Figure</u>		<u>Page</u>
3.7	Structure of the Uniform Quantizer with $L=2^n$ Levels	56
3.8	Performance of the Source Encoded Imagery Transmission System.	74
4.1	Performance of (2,1) Mariner Code.	83
4.2	Performance of (4,3) SPADE Code.	84
4.3	Performance of (8,7) DITEC Code.	84
	Performance of the Source/Channel Encoded Imagery Transmission System:	
4.4	Using (N,K) Block Codes	89
4.5	Using (N,K) Convolutional Codes	90
4.6	With an $m \times m$ Square Block of Low Frequency DCT Coefficients (7,4) Hamming Coded.	93
4.7	With an $m \times m$ Square Block of Low Frequency DCT Coefficients (8,7) DITEC Coded.	95
4.8	With a $t \times t$ Triangular Block of Low Frequency DCT Coefficients (7,4) Hamming Coded.	96
4.9	With a $t \times t$ Triangular Block of Low Frequency DCT Coefficients (8,7) DITEC Coded.	97
4.10	With j Rows and j Columns of Low Frequency DCT Coefficients (7,4) Hamming Coded.	99
4.11	With j Rows and j Columns of Low Frequency DCT Coefficients (8,7) DITEC Coded.	100
4.12	With s of the Most Significant Bits of Each Quantized DCT Coefficient (3,1) Repetition Coded	102
4.13	With s of the Most Significant Bits of Each Quantized DCT Coefficient (2,1) Mariner Coded	103
4.14	Comparison of Source Encoder and Source/Channel Encoder Performance to the Rate Distortion Bound (WMSE)	105

List of Figures (continued)

<u>Figure</u>		<u>Page</u>
4.15	Comparison of Source Encoder and Source/Channel Encoder Performance to the Rate Distortion Bound (MSE)	107
5.1	Flow Chart of Simulated Communication Channel . . .	115
	Comparison of Theoretical and Simulated Source Encoder Performance:	
5.2	For the Truck Image.	118
5.3	For the SMSITE Image	119
5.4	For the Bomber Image	120
5.5	Source Encoded Truck Image Transmitted at P=0 . . .	123
5.6	Source Encoded SMSITE Image Transmitted at P=0. . .	124
5.7	Source Encoded Bomber Image Transmitted at P=0. . .	125
5.8	Source Encoder Performance Under Varying Communication Channel Error Rates	127
5.9	Comparison of Simulated Source Coder Performance for WMSE and MSE Bit Assignment Matrices.	128
5.10	Source Encoded Truck Image Transmitted at P=0 Using MSE Assigned Bit Matrix	129
5.11	Comparison of Source Encoder to Source/Channel Encoder Performance for Simulated Transmissions of Truck Image.	131
5.12	Comparison of Source Coded to Source/Channel Coded Images at P=10	132
5.13	Comparison of Source Encoder to Source/Channel Encoder Performance for Simulated Transmissions of Truck Image; P=10	134
5.14	Source/Channel Encoded Imagery.	136

List of Tables

<u>Table</u>		<u>Page</u>
III.1	Performance of Various Transforms in Encoding a 1-D Gauss-Markov Source	35
III.2	Variances of DCT Coefficients of 2-D Gauss-Markov Image Model	41
III.3	Bit Assignment Matrices for the DCT Coefficients. . .	54
III.4	Optimum Level Spacing and Resultant Mean Square Error for the Uniform Quantizer	59
III.5	Expected Value of the Square of the Uniform Quantizer Output Level.	72
III.6	Performance of the Source Encoder for $P=10^{-3}$	76
IV.1	Summary of Channel Encoder Performance at $P=10^{-3}$. . .	86

Abstract

The performance of an image transmission system which combines an image source encoding technique with communication channel codes is theoretically calculated and measured by simulation. The theoretical and simulated results show that replacing source data bits with error correction coding bits can reduce the distortion in the received image. The magnitude of the reduction depends on the transmission rate, the code rate (N/K) of the channel code, the procedures used to apply the channel code and the error rate of the communication channel. The image source encoding technique used is the two dimensional Discrete Cosine Transform. This technique is combined with both block and convolutional channel codes to form the source/channel encoded imagery transmission system. The distortion in the received image is measured using a weighted mean square error criterion in which the weighting is related to the frequency sensitivity characteristics of the human visual system.

PERFORMANCE OF A SOURCE/CHANNEL ENCODED
IMAGERY TRANSMISSION SYSTEM

I. Introduction

The Air Force is currently involved in several programs which call for the transmission of images or video data. For example, video data links will be required for Remotely Piloted Vehicles and for TV guided bombs. In addition, programs such as the TRITAC Tactical Digital Facsimile call for the transmission of single frames of imagery from reconnaissance sources to strategic and tactical users. In most cases, the communications channels available for imagery transmission are limited in bandwidth and subject to high noise environments, including jamming. These communication channel characteristics are in direct conflict with imagery characteristics.

An image produces a tremendous number of bits of data when converted to a digital format. It is also the nature of imagery that much of this data is redundant. When an image is scanned, as in a television system, one line of the image data is very likely to be similar to the next. Therefore, much research emphasis has been placed on finding source encoding techniques which eliminate this redundancy and reduce the bandwidth requirements for imagery transmission. However, little consideration has been given to the performance of these source encoding techniques in high communication noise environments. Under some of these image compression techniques a single bit error can destroy an entire line or block of image data. Most image compression techniques simply will not operate effectively in high noise environments.

The objective of this thesis is to determine if standard imagery source encoding procedures can be combined with standard channel encoding methods to provide improved imagery transmission performance. That is, if some of

the source data bits are traded off for error correction coding bits, can the overall level of degradation in the received image be reduced?

In order to measure the performance improvement, a criteria for image fidelity must be established. A secondary goal of this thesis is to explore the ramifications of measuring image distortion in terms of weighted mean square error, where the weighting is related to the frequency sensitivity of the human visual system. By using the frequency weighted distortion measure to drive the design of the source/channel encoder, it may be possible to shift the distortion in the image to the spatial frequencies where the eye-brain system has the least sensitivity. In this way, the perceived quality of the received image may be improved.

Background

Research in the field of imagery transmission has taken two paths over the last two decades. The first path has lead to a continuing search for procedures to achieve higher and higher levels of image compression, while maintaining reasonable distortion levels. This line of research is motivated by the need to transmit imagery over narrowband communication channels in a reasonable time frame. It is also motivated by requirements to store imagery in a compact form. The second line of research has concentrated on finding channel encoding procedures which allow error free transmission of imagery over noisy communication links. Much of this research has been conducted in support of space programs such as Mariner and Viking. Channel encoding procedures are also being routinely applied to television signals, which are transmitted over communications satellites. The objective of both lines of research is to achieve performance as close as possible to the theoretical limits.

The theoretical performance limits for data communications are given by information theory. In 1948 Claude E. Shannon laid the foundation of information theory in his classic article "A Mathematical Theory of Communication" (Ref 37). In this paper, Shannon established what is known as the "channel coding theorem". This theorem states that as long as the transmission rate does not exceed the capacity of the communications channel, "appropriate" channel encoding procedures can be found to make the probability of an error occurring in the received data stream as small as desired. This result has lead to a large scale search for the appropriate error correcting codes. The application of channel codes to imagery transmission is only one small part of this much broader research effort.

In 1959, Shannon turned his attention from how to transmit data reliably to what data should be transmitted. In his article "Coding Theorems for a Discrete Source with a Fidelity Criterion" (Ref 38), Shannon established the basic theorems for rate distortion theory. This theory deals with how the data source should be encoded to reduce the redundancy in the transmitted information. Rate distortion theory has been found to be particularly applicable to imagery transmission, due to the high level of redundancy in pictorial information.

Rate distortion theory also provides a method to calculate a bound for source encoder performance. This bound is known as the rate distortion function $R(D)$. The $R(D)$ function is normally interpreted as giving the minimum transmission rate required in order to achieve a desired average distortion level D . For the purposes of this thesis, it will be advantageous to think of this bound as a distortion-rate function. That is, given a fixed transmission rate (R), what is the lowest achievable average distortion (D)? Interpreting the rate distortion function in this manner will allow

a straightforward comparison of the source/channel encoder performance to the theoretical limit.

In order to calculate a rate distortion function, the distortion criteria must be defined. Most previous image compression studies have used mean square error (MSE) as the distortion measure. However, most researchers agree that MSE is not an appropriate measure for the fidelity of imagery. Several authors have postulated that a better measure would be frequency weighted mean square error (WMSE). That is, the contribution of an error to the total measured distortion would depend on the spatial frequency of the image data in which the error occurs. The actual weights in the WMSE distortion measure are related to the frequency response characteristics of the human visual system. The spatial frequencies which the eye-brain system is the most sensitive to are weighted heavily. By using the WMSE distortion criteria in the design of the source channel encoder, the frequencies that carry the greatest weight will be accurately reproduced, thus providing an improved quality image.

Organization of the Thesis

The research for this thesis was conducted in two phases. First, a theoretical analysis was performed based on a statistical image model to design the source/channel encoder and calculate its performance. Then, the source/channel encoder was simulated on a digital computer. The simulated system was used to "transmit" real images and measure the distortion in the received image.

In order to establish a standard to judge the performance of the source/channel encoder, the rate distortion function for an image was calculated. This required the development of a statistical model for an image and a

definition of the distortion measure. The research focused on transmitting single frames of imagery, since this covers all requirements if video is considered to be a series of individual images. The statistical model for an individual image and the distortion measure are described in Chapter II. The calculation procedures for the rate distortion function and the calculated $R(D)$ curves are also presented in Chapter II.

Since the primary objective of the research was to determine the effect of adding channel encoding to the imagery transmission system, only one source encoding technique was used. The technique selected is the two dimensional discrete cosine transform (DCT). Chapter III gives the reasons for selecting the DCT source encoder, describes how it is implemented and presents a theoretical performance calculation for it.

Chapter IV presents the theoretical results of trading off source data bits for error correction coding bits. The chapter begins by describing the channel codes used in the research. Both block and convolutional codes were investigated. Then, the procedures for applying the channel codes are discussed. Finally, a number of graphs are presented to show the performance improvement available through the application of channel codes.

In order to confirm the theoretical calculations, an image transmission simulation was performed. The simulation is the subject of Chapter V. The hardware and software used in the simulation are described, and the simulated results are presented in a set of images and graphs.

Finally, Chapter VI summarizes the theoretical calculations and simulated results and gives recommendations for further work.

II. Rate Distortion Curve for an Imagery Source

The rate distortion function, $R(D)$, gives the minimum rate (R) required to represent a source symbol if an average distortion (D) is allowable (Ref 25:73). The rate distortion function will be used as a standard against which the performance of the imagery source/channel coding methods will be compared. In order to calculate an $R(D)$ curve the statistics of the data source must be known. Also, the functional form of the distortion measure (D) must be given. These two topics are examined in the next two sections of this paper.

Image Model

The primary concern of this thesis is the transmission of "still" or single frame monochrome images. Further, it will be assumed that the image has been converted to a digital form using sampling procedures which minimize any distortion effects due to aliasing. Therefore the "source" or original input to the coding procedure is considered to be a set of digitized samples of an analog image. These samples are commonly referred to as pixels.

Image data is inherently two dimensional. For simplicity of calculation (and for compatibility with the image processing hardware used in the image transmission simulation discussed in Chapter V), the original image will be defined to be an $M \times M$ square array of pixels where M is a power of 2. This definition in no way limits the applicability of the results derived.

Statistical Properties of Imagery. While the form of the original image is easily defined, the content or statistical properties of it are

not so easily determined. Histograms of digitized real world images vary greatly from one another and do not appear to consistently fit a known probability distribution. As an example of this, several of the original images used in the simulation along with their histograms are presented in Figures 2.1-2.3. Several authors have noted one common trait shared by most images (Ref 17;32;39). There is normally a significant amount of correlation between adjacent pixels of real pictures. For this reason images are often modeled as two dimensional auto-regressive fields (Ref 16;26;39).

That is, if $f(i,j)$ represents the intensity or gray scale value of the (i,j) pixel, then

$$f(i,j) = \sum_{k=0}^K \sum_{l=0}^L' a(k,l) f(i-k, j-l) + W(i,j) \quad (2.1)$$

where the prime ('') indicates that the $k=l=0$ point is not included in the sum and where $W(i,j)$ is a two dimensional, zero mean sequence of independently, identically distributed random variables. In order to completely characterize the image, the values of K , L and the weights $a(k,l)$ must be known and the distribution of the $W(i,j)$ must be given.

Again, turning to existing research literature, several authors have shown that imagery can be successfully modeled as a two dimensional homogeneous Markov process (Ref 16;17;28;39). That is, K and L can both be set equal to 1 and the $a(k,l)$ can be defined in terms of the row and column correlation coefficients r_R and r_C .

$$f(i,j) = r_R f(i-1,j) + r_C f(i,j-1) - r_R r_C f(i-1,j-1) + W(i,j) \quad (2.2)$$

The final ingredient needed to define the image statistics is the

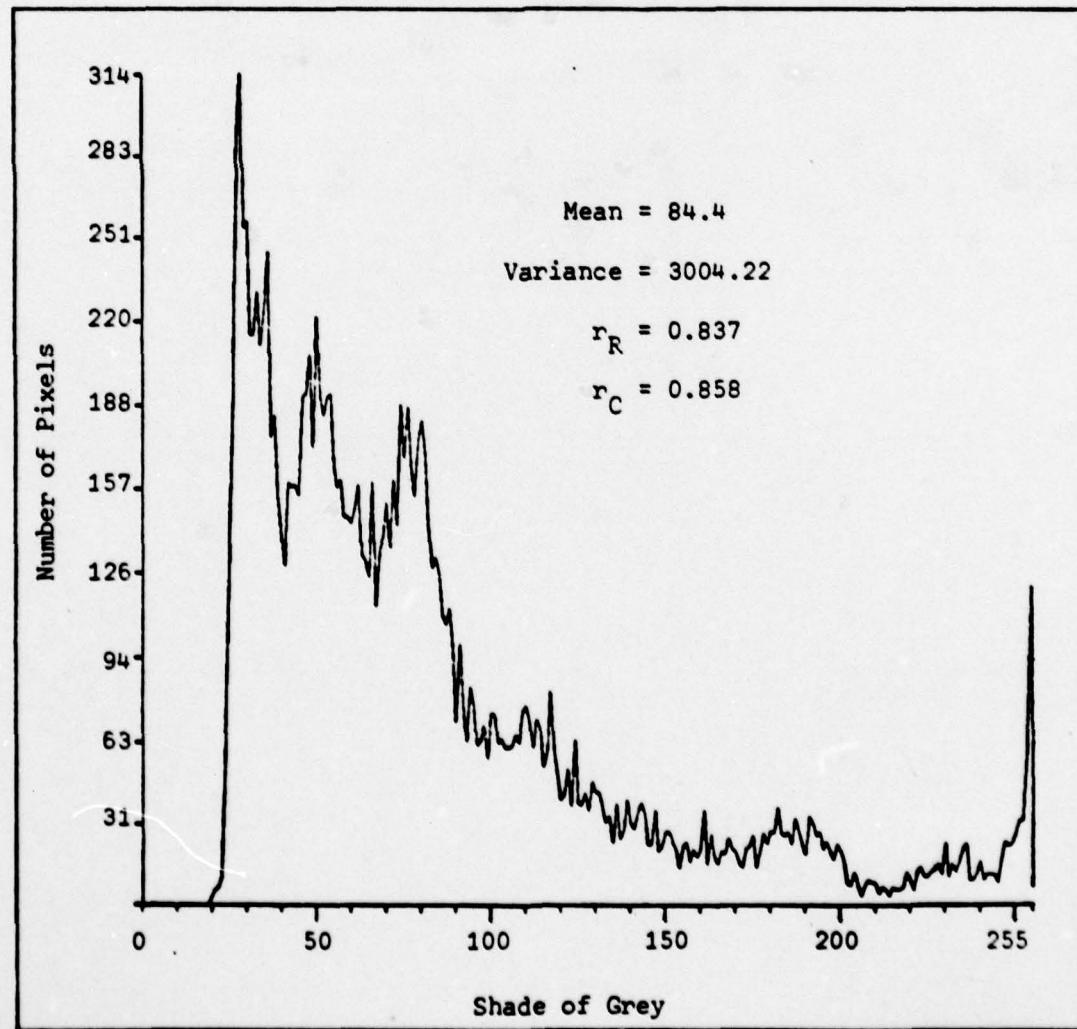


Figure 2.1 Truck: Original Image and Histogram

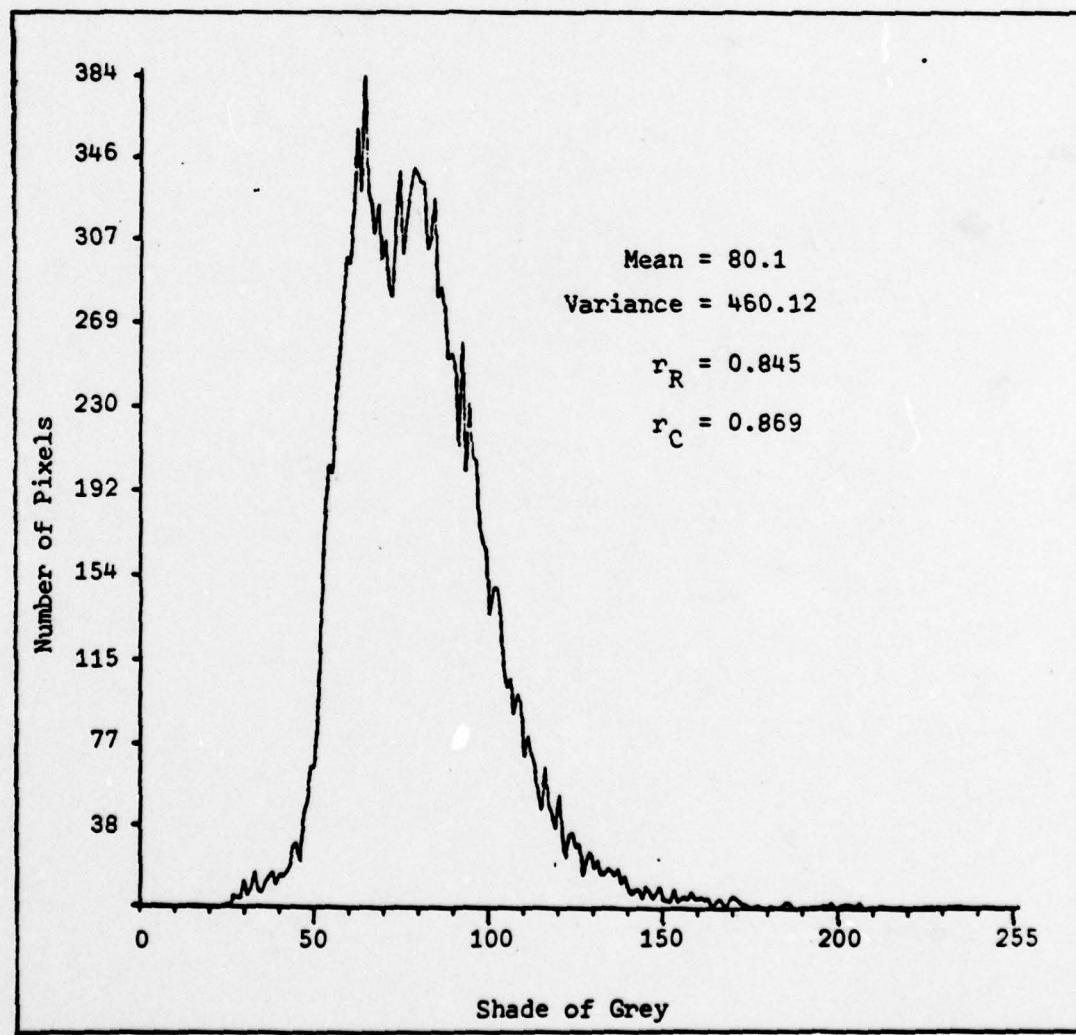
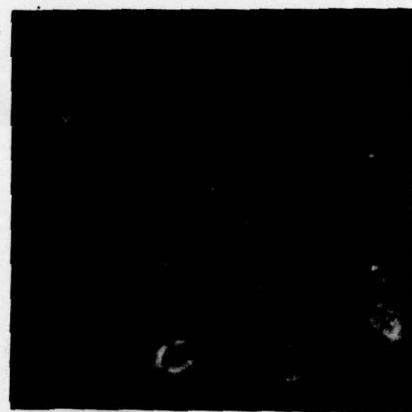


Figure 2.2 SMSITE: Original Image and Histogram

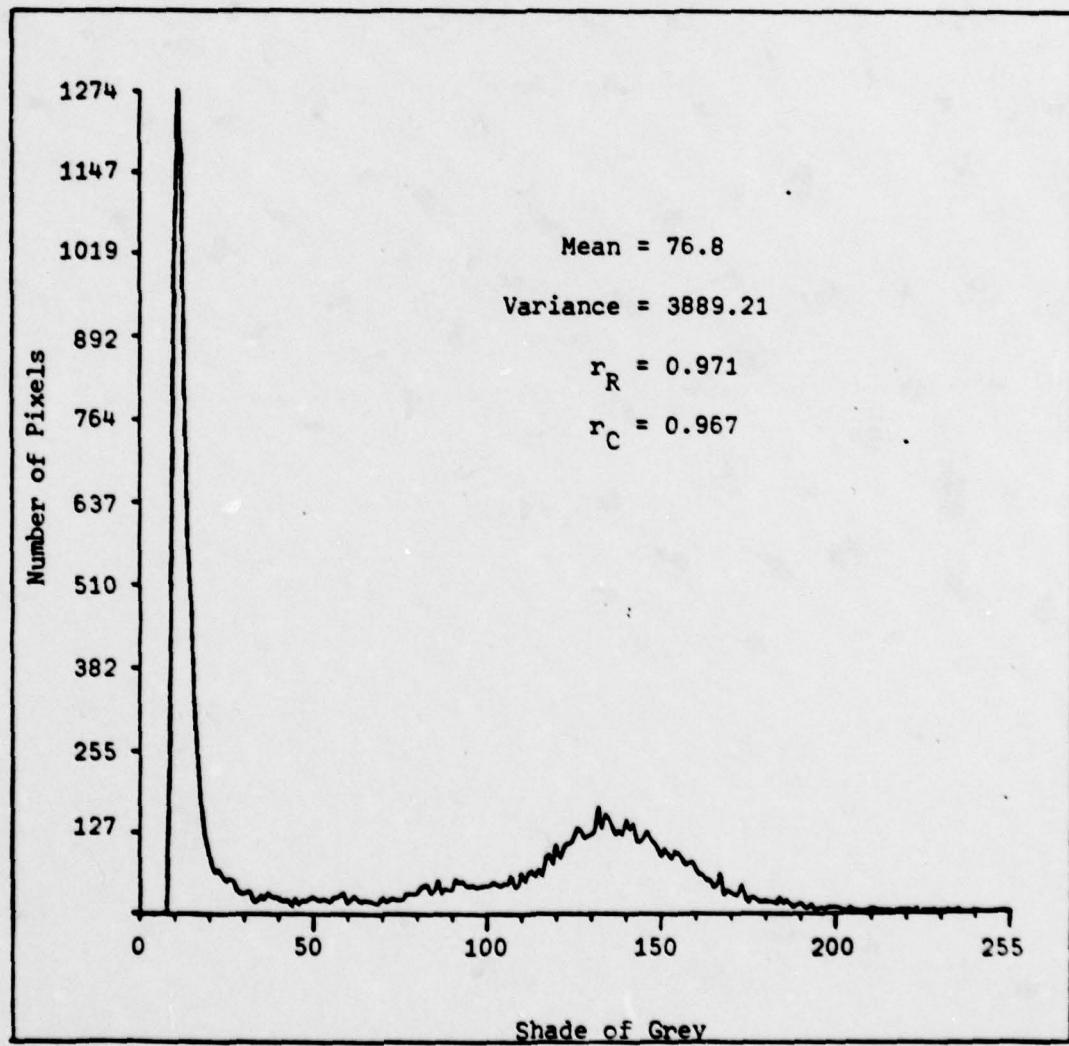


Figure 2.3 Bomber: Original Image and Histogram

distribution of the error terms, $W(i,j)$. At this point, it is necessary to take a look forward to determine the intended use of the image model. The first thing seen is a calculation of a rate distortion function. The calculation of $R(D)$ for sources with memory is extremely difficult and in fact has only been carried out for Gaussian sources. Therefore in order to continue with the analysis an assumption must be made that the $W(i,j)$ are Gaussian.

On the surface, this assumption appears to have little support other than analytical tractability, since in general real world images are not Gaussian. Sakrison and Algazi (Ref 36:387) have noted that this assumption is useful in that it results in the calculation of the "worst case" rate distortion curve. Consider a class A of sources (homogeneous random fields with identical power spectral densities). The rate distortion function of the class $R_A(D)$ is the minimum rate required by any single encoder that guarantees that the average distortion for each probability distribution p in the class will be less than or equal to D . That is

$$R_A(D) \geq \sup_{p \in A} R_p(D) \quad (2.3)$$

where $R_p(D)$ is the rate distortion function for the source with distribution p . For the squared error distortion measure it can be shown that $R_A(D) = R_G(D)$ where $R_G(D)$ is the rate distortion function of a Gaussian random field. Therefore, assuming that the $W(i,j)$ are Gaussian distributed guarantees that the actual level of distortion (MSE) achievable for any fixed value of R is less than or equal to the calculated value.

Correlation and Power Spectral Density of the Image Model. Under

the assumption that the image data comes from a homogeneous Gauss-Markov Process, the auto correlation function, $R(i,j)$, is given by

$$R(i,j) = \sigma_f^2 r_R^{|i|} r_C^{|j|} \quad (2.4)$$

where σ_f^2 is the common variance of the $f(j,k)$. The corresponding power spectral density can be shown to be (Ref 5:113)

$$S(w_1, w_2) = \frac{\sigma_f^2 (1-r_R^2)(1-r_C^2)}{(1-2r_R \cos w_1 + r_R^2)(1-2r_C \cos w_2 + r_C^2)} \quad (2.5)$$

Both of these equations assume that the mean of the image samples is zero. This can be accomplished by determining a sample mean of the $f(i,j)$ and subtracting it from the pixel values. Figure 2.4 shows a perspective plot of the power spectral density of the image model.

Distortion Measure

In addition to a statistical model for the imagery, the other element needed to calculate a rate distortion function is the error metric. The distortion measure most often used in image coding studies is mean square error (MSE).

$$MSE = E \left\{ \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} [f(i,j) - g(i,j)]^2 \right\} \quad (2.6)$$

where

$f(i,j)$ = original pixel values prior to transmission

$g(i,j)$ = received and reconstructed pixel values

This quantity is often normalized by the total energy in the original image.

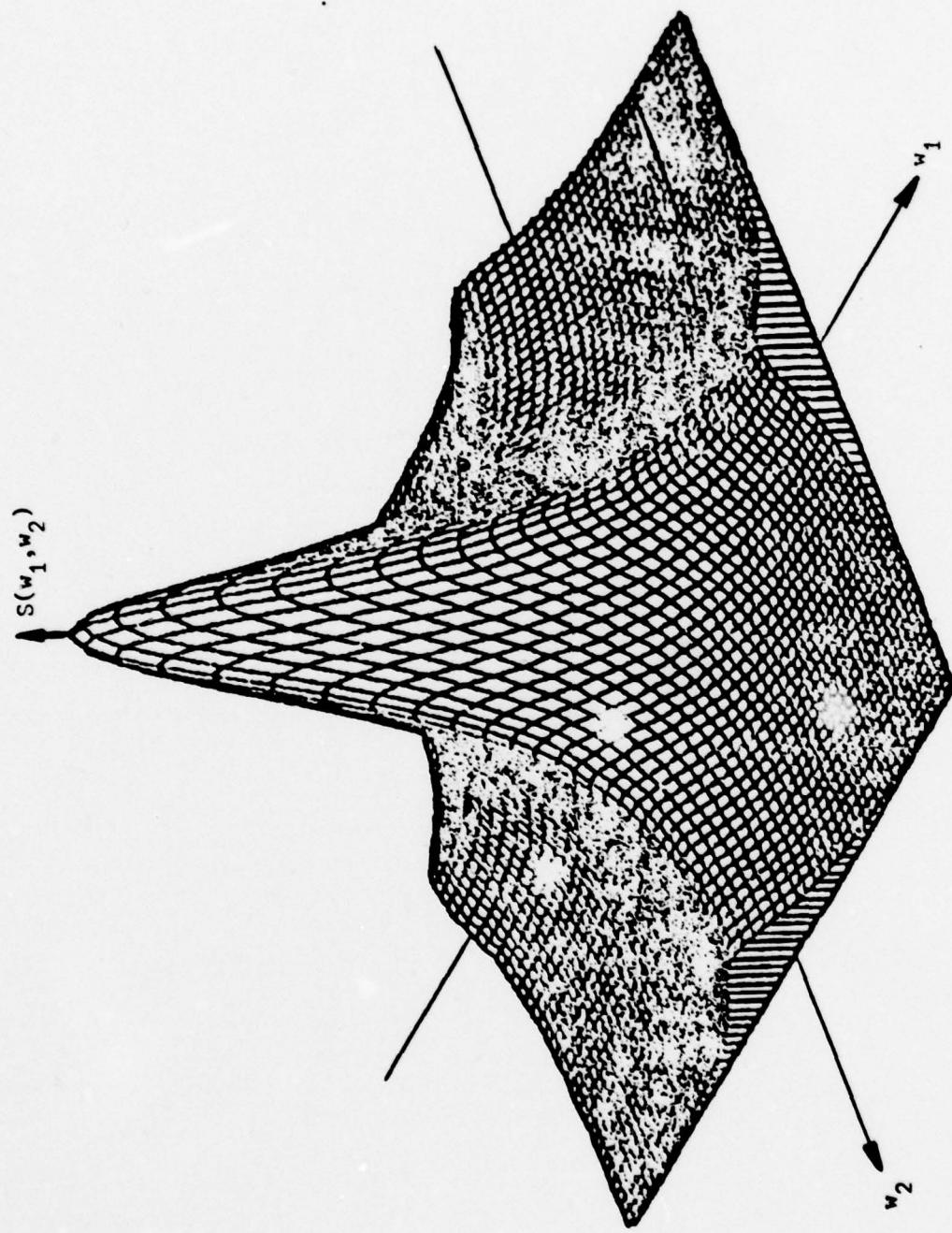


Figure 2.4 Power Spectral Density of Gauss-Markov Image Model ($r_R = r_C = 0.9$)

$$NMSE = \frac{E \left\{ \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} [f(i,j) - g(i,j)]^2 \right\}}{E \left\{ \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} [f(i,j)]^2 \right\}} \quad (2.7)$$

The beauty of the MSE distortion measure lies in its analytic tractability and in the extensive amount of work which has been accomplished in calculating R(D) for it. However, MSE is not a good distortion measure for image studies because it is a global measure and gives no information on the spatial structure or frequency distribution of the errors within the image (Ref 40:31). The human eye-brain system is known to be very sensitive to image frequencies in the range of 4 to 10 cycles/degree of visual field at normal photopic levels of brightness. It is very insensitive to frequencies above 30 cycles/degree (Ref 6;12;41).

In order to account for the characteristics of human vision the distortion measure used in this report will be a weighted mean square error criteria.

$$WMSE = E \left\{ \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} [a(i,j) * \{f(i,j) - g(i,j)\}]^2 \right\} \quad (2.8)$$

where $a(i,j)$ is the weighting function in the spatial domain and the asterisk indicates convolution.

The actual form of the weighting function has been the subject of several papers (Andrews and Hall; Hall and Hall; Mannos and Sakrison; Ref 3;18;23). In general these authors have postulated weighting functions based on the experimental measurement of the contrast sensitivity of the human visual system conducted by a number of researchers (Campbell and Robson; VanNess and Bouman; Dipalma and Lowry; Ref 6;12;41).

The experimental evidence shows that the eye has a distinctive

contrast threshold which varies with the frequency of the image information. Figure 2.5 gives an example of the measured contrast threshold function, $C(f)$. Campbell and Robson also found some evidence to indicate that the human visual system could be modeled as a set of linearly operating independent mechanisms selectively sensitive to limited ranges of spatial frequencies (Ref 6:551). On the basis of this evidence, the human visual system can be treated as a band-pass filter with a transfer function $A(f)$ proportional to $1/C(f)$.

Two $A(f)$ functions in particular were investigated for application as the weighting criteria in the WMSE distortion measure. Both functions assume that the contrast sensitivity of the eye is radially symmetric. Therefore they are expressed as a function of radial frequency, f_r , in units of cycles/degree of visual field. The first of these functions was proposed by Mannos and Sakrison and is given by (Ref 23:79)

$$A_S(f_r) = 2.6(0.0192 + 0.0114f_r) \exp \left\{ -(0.114f_r)^{1.1} \right\} \quad (2.9)$$

The second weighting function discussed by Ratliff as well as others (Ref 33) has the following form

$$A_R(f_r) = K_1 \exp \{ -(f_r \sigma_1)^2 \} - K_2 \exp \{ -(f_r \sigma_2)^2 \} \quad (2.10)$$

where the values of K_1 , K_2 and σ_1 and σ_2 depend on the ambient lighting conditions. Carl has studied the functional dependence of these parameters on luminance and found (Ref 7)

$$K_1 = 272.516 + 70.362 \log_{10} (B) \quad (2.11)$$

$$K_2 = 225.900 + 65.4585 \log_{10} (B)$$

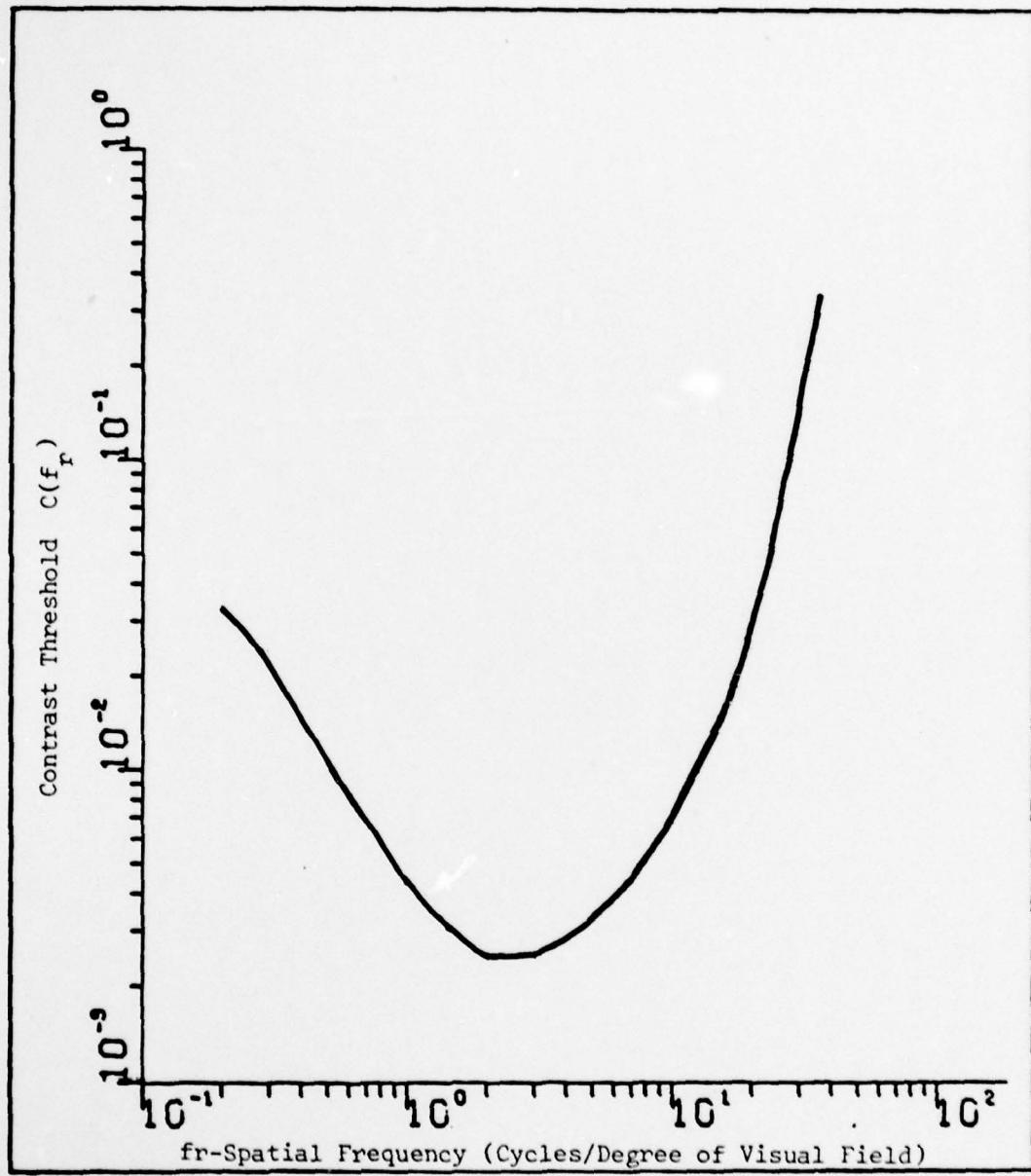


Figure 2.5 Measured Contrast Threshold of the Human Visual System
(Ref 6:554)

$$\sigma_1 = \begin{cases} 0.1279 B^{-0.1147} & , \quad B \leq 2500 \\ 0.0521302 & , \quad B > 2500 \end{cases} \quad (2.11)$$

$$\sigma_2 = \begin{cases} \sigma_1 & , \quad B < 1 \\ 0.685 - 0.557 B^{-0.161} & , \quad 1 < B \leq 2500 \\ 0.526671 & , \quad B > 2500 \end{cases}$$

where B = average luminance in Trolands.

The two $A(f)$ functions, normalized to have a peak value of 1, are plotted in Figure 2.6. The Sakrison function is based on data taken at an average luminance of 2000 Trolands. The corresponding curve for the Ratliff model can be seen to closely approximate the Sakrison curve. Both functions peak in the 4 to 10 cycles/degree range and fall off sharply above 30 cycles/degree. Because of the close correspondence between the two curves only the Ratliff model was used throughout the remainder of the thesis research and will be designated simply by $A(f_r)$ henceforth.

In calculating the rate distortion function the weighting curve must be expressed as a function of w_1 and w_2 , the two dimensional frequency components of the image power spectral density. These frequency terms have units of radians per pixel width while f_r has units of cycles/degree of visual field. To convert to the new units the conditions under which the image will be viewed must be known.

The industry standard viewing conditions for images on graphics and television displays are a twelve inch square picture viewed at a distance of thirty inches with a resolution of no less than sixty lines per inch (Ref 13:S3). After some basic geometrical considerations this translates to thirty lines of resolution (or pixel widths) per visual

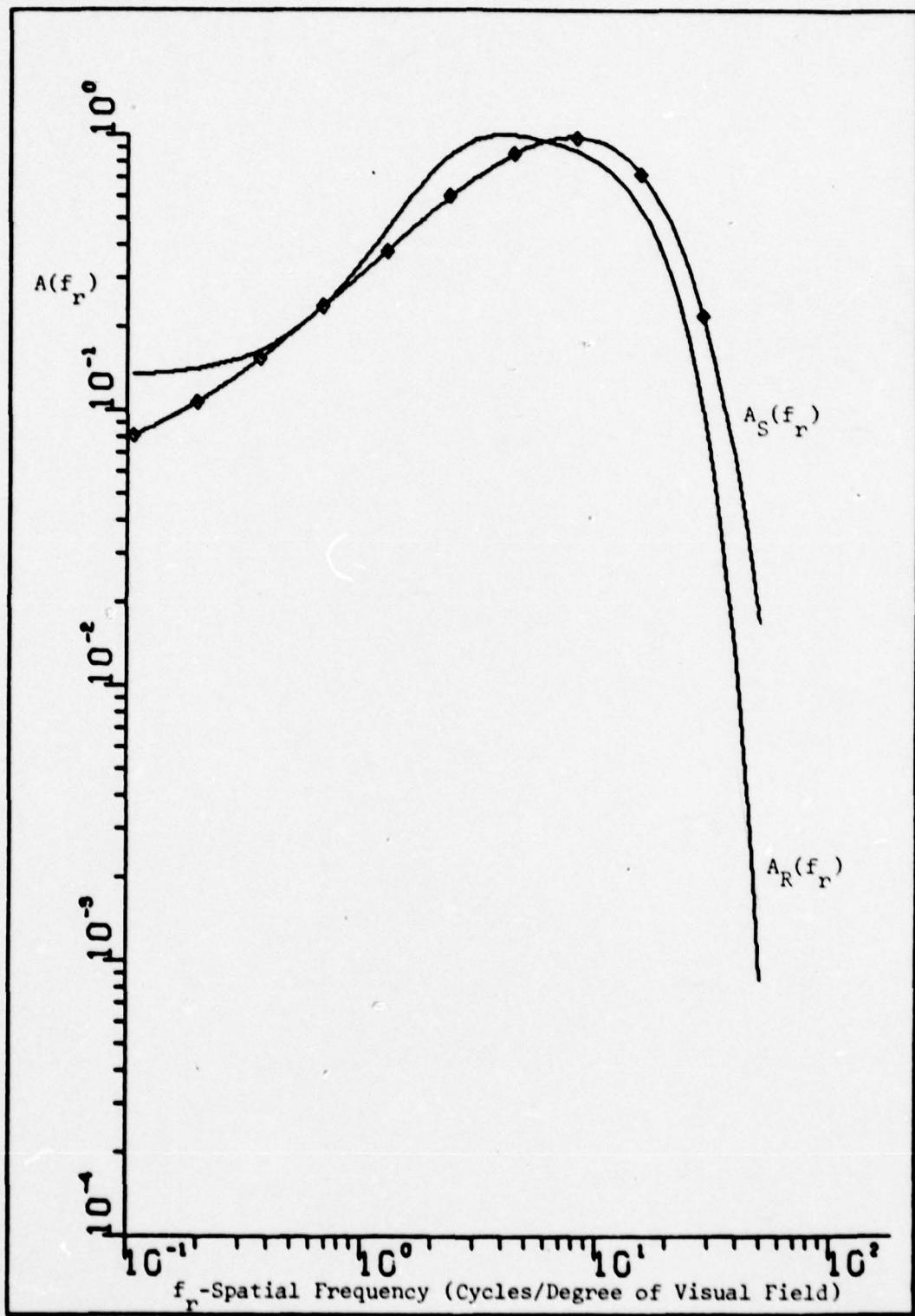


Figure 2.6 Models for Human Visual System Transfer Function

degree. Therefore

$$f_r \left[\frac{\text{cycles}}{\text{deg of vision}} \right] = w_r \left[\frac{\text{radians}}{\text{pixel}} \right] \times \frac{1}{2\pi} \left[\frac{\text{cycles}}{\text{radian}} \right] \times 30 \left[\frac{\text{pixels}}{\text{deg of vision}} \right] \quad (2.12a)$$

and

$$w_r = \sqrt{w_1^2 + w_2^2} \quad (2.12b)$$

So

$$f_r = \frac{\sqrt{w_1^2 + w_2^2}}{\pi} \cdot 15 \quad (2.12c)$$

Substitution of (2.12) into (2.10) gives

$$A(w_1, w_2) = K_1 \exp \left\{ \frac{-(w_1^2 + w_2^2)225 \sigma_1^2}{\pi^2} \right\} \quad (2.13)$$

$$-K_2 \exp \left\{ \frac{-(w_1^2 + w_2^2)225 \sigma_2^2}{\pi^2} \right\}$$

Image Rate Distortion Function

Rate Distortion Function for MSE Distortion Criteria. The rate distortion function of a Gaussian process with a mean square error distortion criteria has been thoroughly investigated. For a memoryless Gaussian source with arbitrary mean and variance σ^2 , the $R(D)$ function is given by the well known equation (Ref 5:99)

$$R(D) = \begin{cases} \frac{1}{2} \log \frac{\sigma^2}{D} , & 0 \leq D \leq \sigma^2 \\ 0 , & D \geq \sigma^2 \end{cases} \quad (2.14)$$

Berger has shown that for a one dimensional auto-regressive Gaussian source with power spectral density $S(w)$ the rate distortion function for the MSE distortion criteria is given parametrically by (Ref 5:112)

$$D(Q) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \min [Q, S(w)] dw \quad (2.15)$$

$$R[D(Q)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \max \left[0, \frac{1}{2} \log \frac{S(w)}{Q} \right] dw$$

The procedure for calculating $D(Q)$ is illustrated in Figure 2.7 for the one dimensional Gauss-Markov case in which $S(w) = 1-r^2/1-2rcosw+r^2$. $D(Q)$ is calculated by integrating the area under the bold line from $-\pi$ to π . Similarly, $R[D(Q)]$ is calculated by integrating $1/2\log\{S(w)/Q\}$ from $-w_0$ to w_0 .

Stuller and Kurz expanded Berger's analysis to show that for a two dimensional auto-regressive Gaussian source (such as that given by (2.1)) the MSE rate distortion function becomes (Ref 39:490)

$$D(Q) = \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \min [Q, S(w_1, w_2)] dw_1 dw_2 \quad (2.16)$$

$$R[D(Q)] = \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \max \left\{ 0, \frac{1}{2} \log \left[\frac{S(w_1, w_2)}{Q} \right] \right\} dw_1 dw_2$$

where $S(w_1, w_2)$ is the two dimensional power spectral density of the source process.

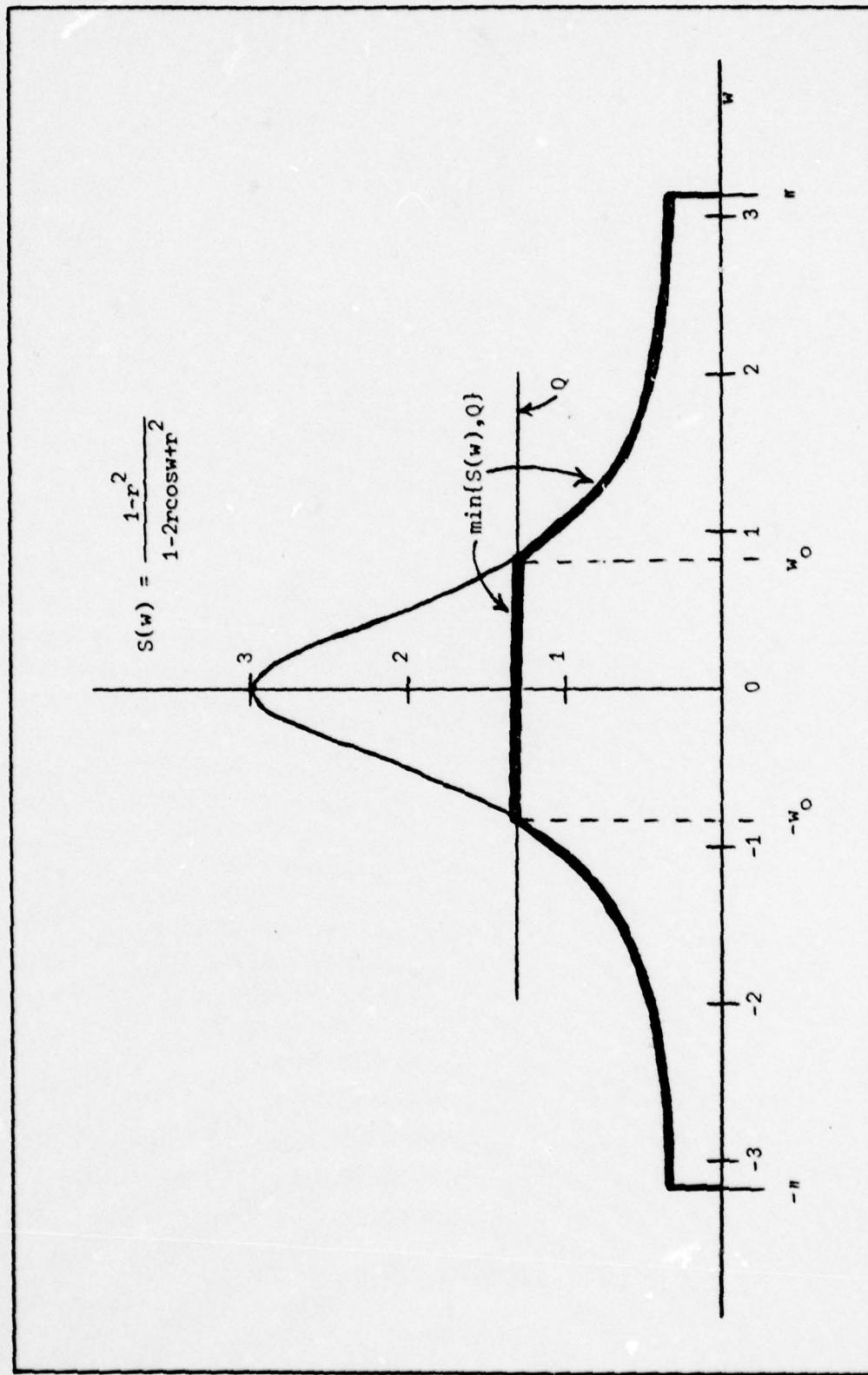


Figure 2.7 Calculation Procedure for Mean Square Error Rate Distortion Function
(1-D Gauss-Markov Source $r=0.5$)

Rate Distortion Function for Weighted MSE Criteria. The rate distortion function of a Gaussian source for the weighted mean square error criterion defined by (2.8) turns out to be very similar in form to the unweighted version. According to the development given by Sakrison and Algazi the parametric representation of the rate distortion function for WMSE is (Ref 36:390)

$$D(Q) = \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \min \left[Q, A^2(w_1, w_2) S(w_1, w_2) \right] dw_1 dw_2 \quad (2.17)$$

$$R[D(Q)] = \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \max \left\{ 0, \frac{1}{2} \log \left[\frac{A^2(w_1, w_2) S(w_1, w_2)}{Q} \right] \right\} dw_1 dw_2$$

where $A(w_1, w_2)$ is the Fourier transform of the function $a(i, j)$ in equation (2.8). For the weighting function used in this report, $A(w_1, w_2)$ was determined experimentally and is defined by equation (2.13).

Comparing equations (2.17) and (2.16) it can be seen that the $R(D)$ function for the WMSE case is equivalent to calculating an MSE rate distortion function for a process with power spectral density $A^2(w_1, w_2) S(w_1, w_2)$. By rearranging the terms in (2.17) slightly, a more physical interpretation can be developed. Defining $Q' = Q/A^2(w_1, w_2)$, the equations of (2.17) become

$$D(Q') = \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \min \left[A^2(w_1, w_2) Q', A^2(w_1, w_2) S(w_1, w_2) \right] dw_1 dw_2 \quad (2.18)$$

$$R[D(Q')] = \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \max \left\{ 0, \frac{1}{2} \log \left[\frac{S(w_1, w_2)}{Q'} \right] \right\} dw_1 dw_2$$

Since $A^2(w_1, w_2) \geq 0$, the distortion equation can be rewritten as

$$D(Q') = \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} A^2(w_1, w_2) \min\left[Q', S(w_1, w_2)\right] dw_1 dw_2 \quad (2.19)$$

The distortion is now clearly a frequency weighted error.

A comparison of the rate equation of (2.18) with the corresponding MSE version given in (2.16) reveals that the two are identical in structure. The only difference is the threshold Q' is a function of the frequency. In fact Q' can be shown to be a very specific frequency function. Q' is defined in terms of $A(w_1, w_2)$ which was derived by a change in units from $A(f)$. In turn $A(f)$ was proposed to be proportional to the inverse of the contrast threshold function, $C(f)$, of the human visual system. Therefore, the threshold Q' is given by

$$Q' = K^2 C^2(f) \quad (2.20)$$

where K is a proportionality constant. This implies that the rate of the source for the weighted MSE criteria is calculated by forming the ratio of the power spectral density of the source process to the square of the contrast threshold. Figure 2.8 gives a one dimensional view of this calculation procedure.

Numerical Calculation of the Image Rate Distortion Function. Due to the form of the weighting function and power spectral density in equation (2.17) no effort was made to find a closed form solution of the rate distortion equation. Instead the integrals were calculated numerically. The results are presented in graphical form in Figures 2.9

R found by integrating
 $\frac{1}{2} \log_2 x/y$ from $-w_0$ to w_0

D found by multiplying $A^2(w)$
 times function given by bold
 line and integrating from
 $-\pi$ to π

$$S(w) = \frac{1-r^2}{1-2r \cos w + r^2}$$

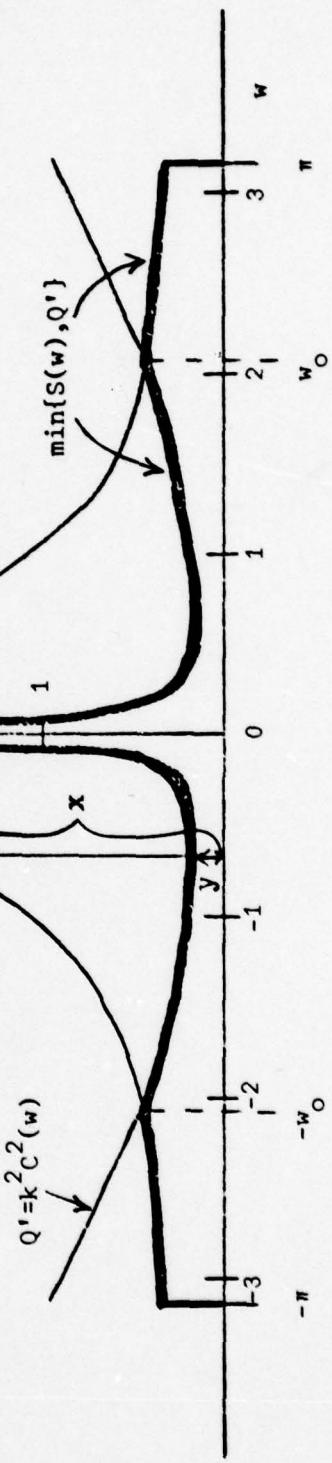


Figure 2.8 Calculation Procedure for Weighted Mean Square Error Rate Distortion Function
 (1-D Gauss-Markov Source $r=0.5$)

through 2.12.

The first set of curves (Figure 2.9) gives the mean square error $R(D)$ function of a two dimensional Gauss-Markov source with $\sigma_f^2 = 1$.

The row and column correlation coefficients are different for each curve. Because the process is assumed to have unit variance (i.e. the total average power of the source is equal to 1) the maximum value that the distortion can take on is also 1.

$$D_{\max} = \sigma_f^2 = \left(\frac{1}{2\pi}\right)^2 \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} S(w_1, w_2) dw_1 dw_2 = 1 \quad (2.21)$$

Figure 2.9 shows that as the correlation between picture elements decreases the rate required to achieve a fixed distortion level increases.

A comparison of weighted and unweighted mean square error $R(D)$ functions is presented in Figure 2.10. The weighted distortion curve is presented in two forms. First the $R(D)$ function was calculated by normalizing the weighting function $A(w_1, w_2)$ so the total weighted power in the signal was equal to 1.

$$\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} A_p^2(w_1, w_2) S(w_1, w_2) dw_1 dw_2 = 1 \quad (2.22)$$

$$A_p^2(w_1, w_2) = \frac{A^2(w_1, w_2)}{\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} A^2(w_1, w_2) S(w_1, w_2) dw_1 dw_2} \quad (2.23)$$

The rate distortion equations of (2.17) were then calculated using this power normalized weighting function, $A_p(w_1, w_2)$. The resulting curve lies entirely above the mean square error curve. This implies that for any fixed level of distortion more bits are required to represent the weighted

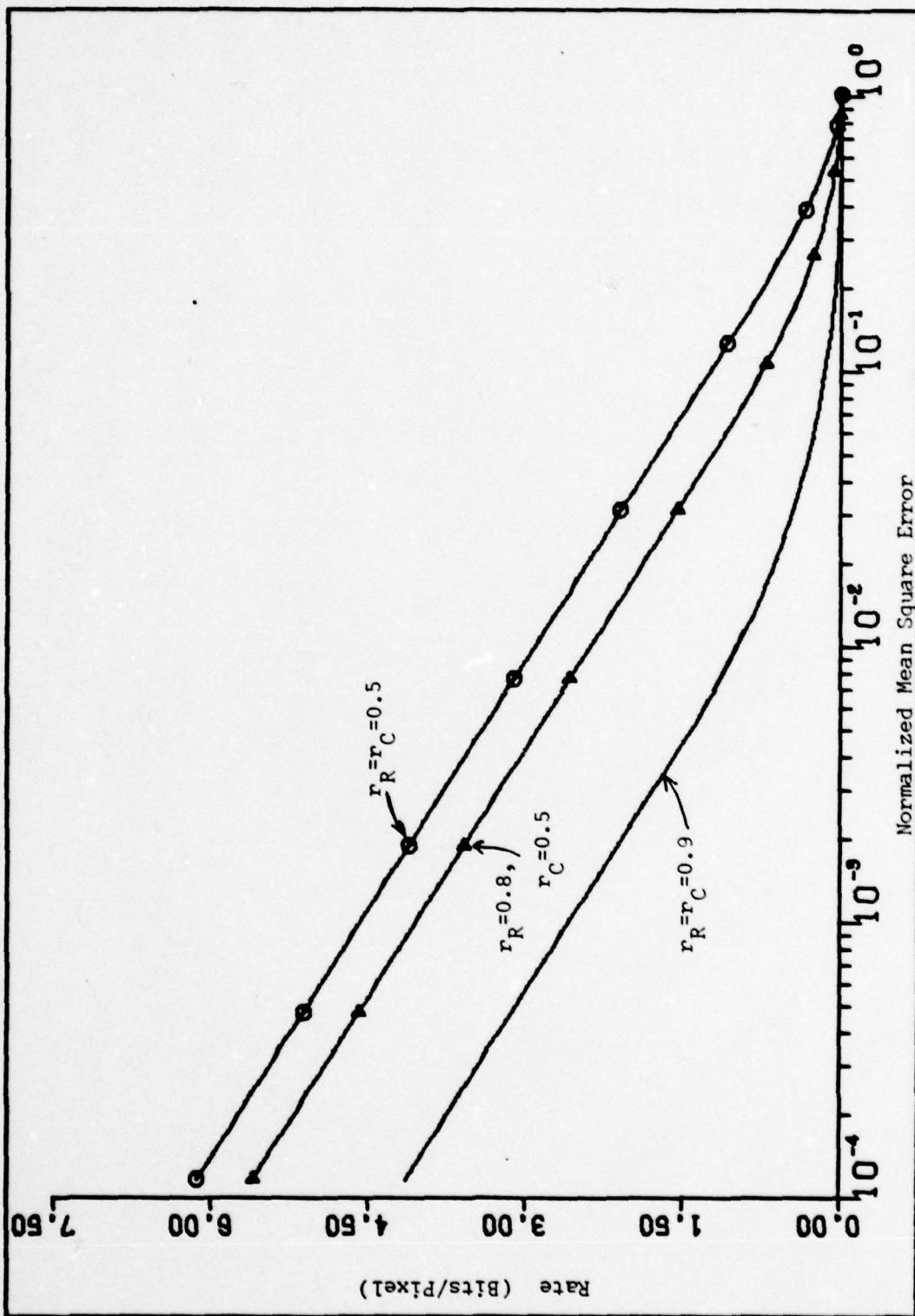


Figure 2.9 Rate Distortion Function for the 2-D Gauss-Markov Image Model

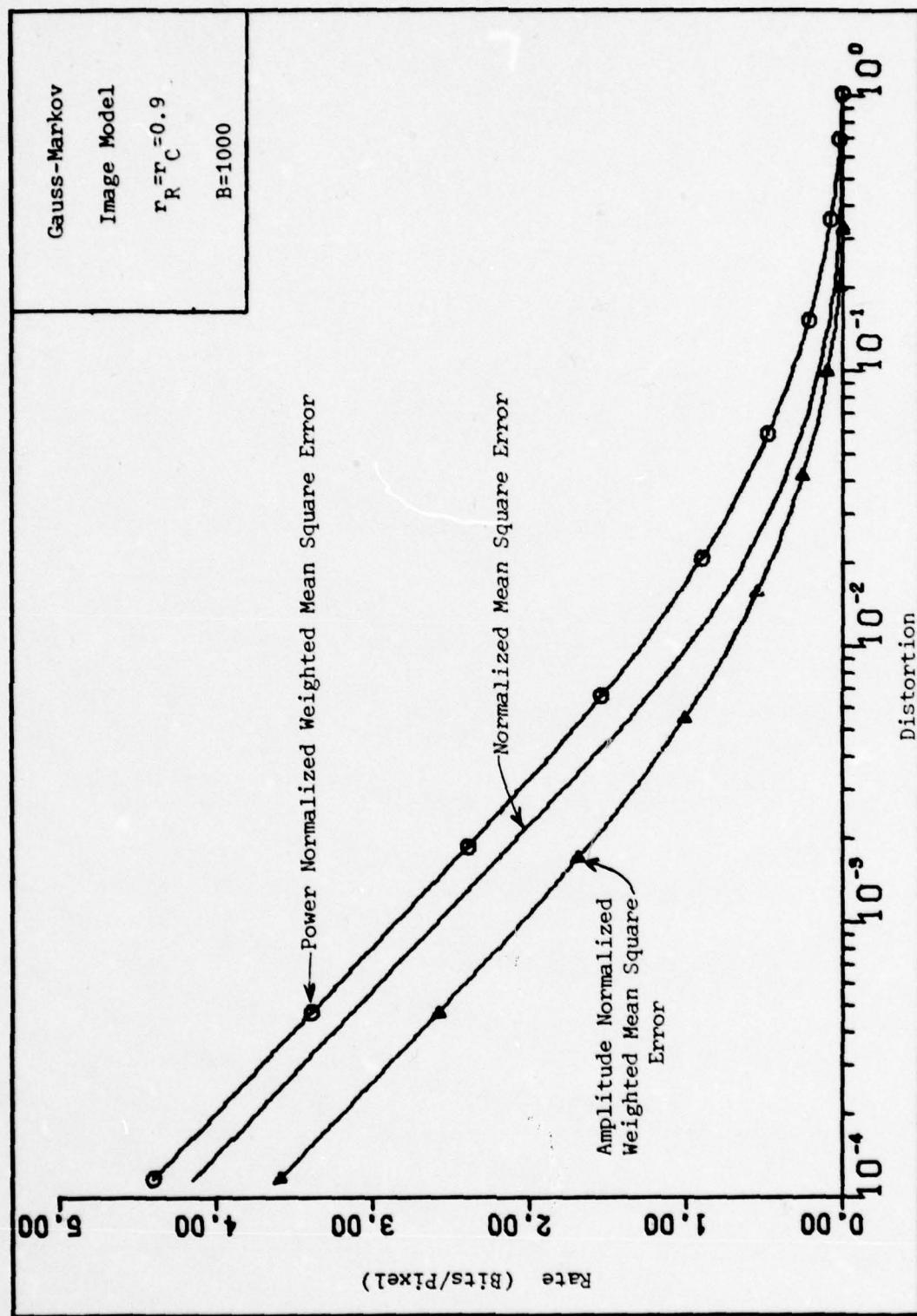


Figure 2.10 Comparison of WMSE and MSE Rate Distortion Functions

unity power source than the unweighted unity power source. Conversely, given a fixed rate the distortion, as measured in a weighted sense, will exceed that for the unweighted case. The reason for this is the strong accentuation of the mid-frequency ranges provided by the weighting function.

Figure 2.10 also presents a second view of the frequency weighted $R(D)$ function. If instead of using the weighting function to accentuate the mid-frequencies, we use it to deemphasize the low and high frequency content, which the eye is less sensitive to, we may be able to save transmission bits. This effect is demonstrated by normalizing $A(w_1, w_2)$ such that its peak magnitude (which occurs in the mid-frequency band) is set equal to 1.

$$A_n(w_1, w_2) = \frac{A(w_1, w_2)}{\text{MAX } A(w_1, w_2)} \quad (2.24)$$

Figure 2.10 shows that using this amplitude normalized weighting function does result in the corresponding $R(D)$ curve falling below the MSE result. This too is to be expected since the total power in the $A_n(w_1, w_2)$ weighted source is less than that for the unweighted source.

The final two sets of rate distortion curves show the effect of variation in the average luminance (B) under which the image is to be viewed. Since the frequency weighting function depends on B some variation in the $R(D)$ functions is to be expected. Figures 2.11 and 2.12 show that the sensitivity of the $R(D)$ function to variations in B is small for values of B in the range of 300 to 2000 Trolands. This range covers most viewing conditions from normal office level lighting to very bright illumination such as that needed for photo interpretation. Since the sensitivity of $R(D)$ with respect to variations of B is small one value of B (1000 Trolands) is used in all further calculations.

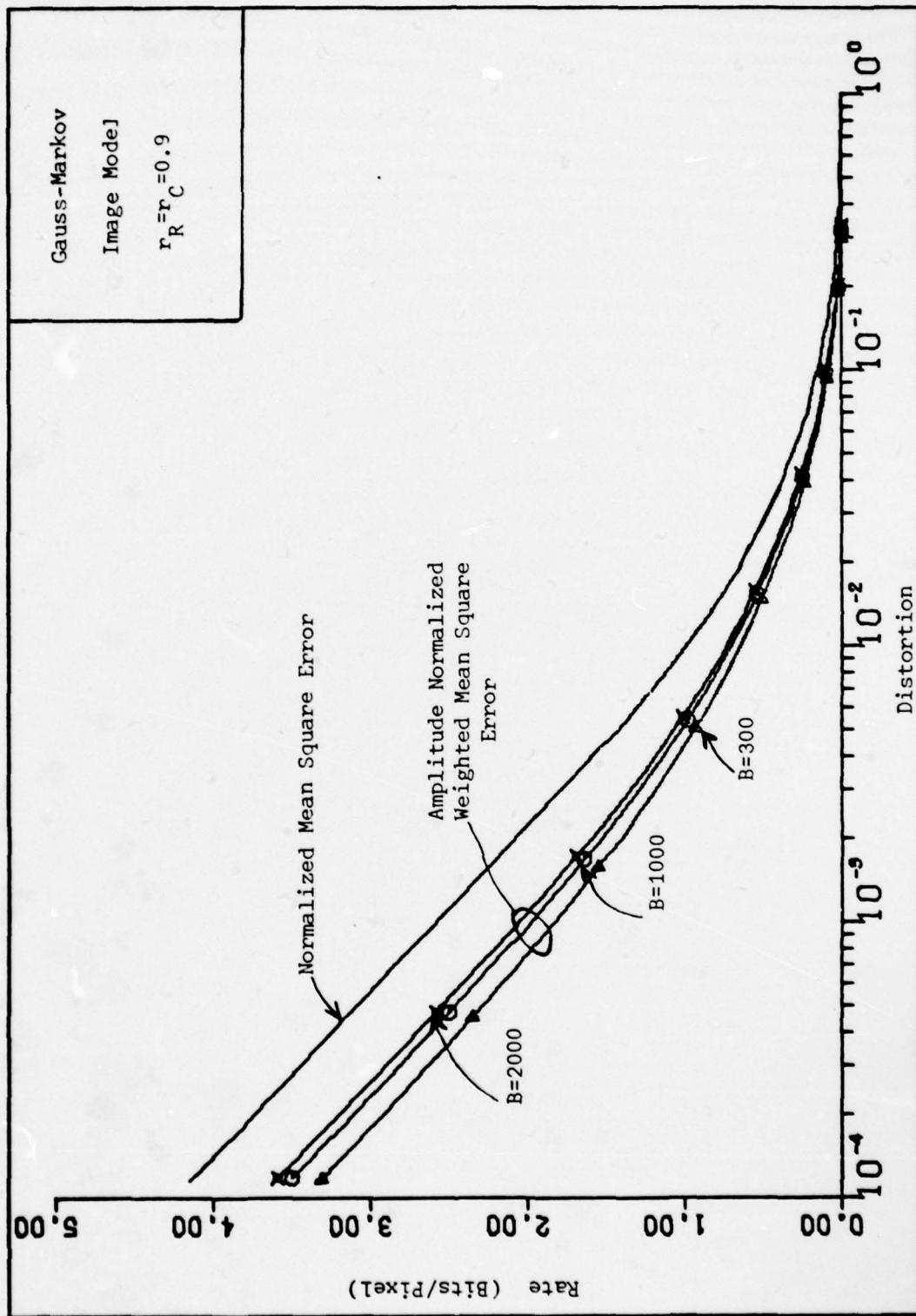


Figure 2.11 Dependence of $R(D)$ on Average Illumination (B) for Amplitude Normalized WMSE

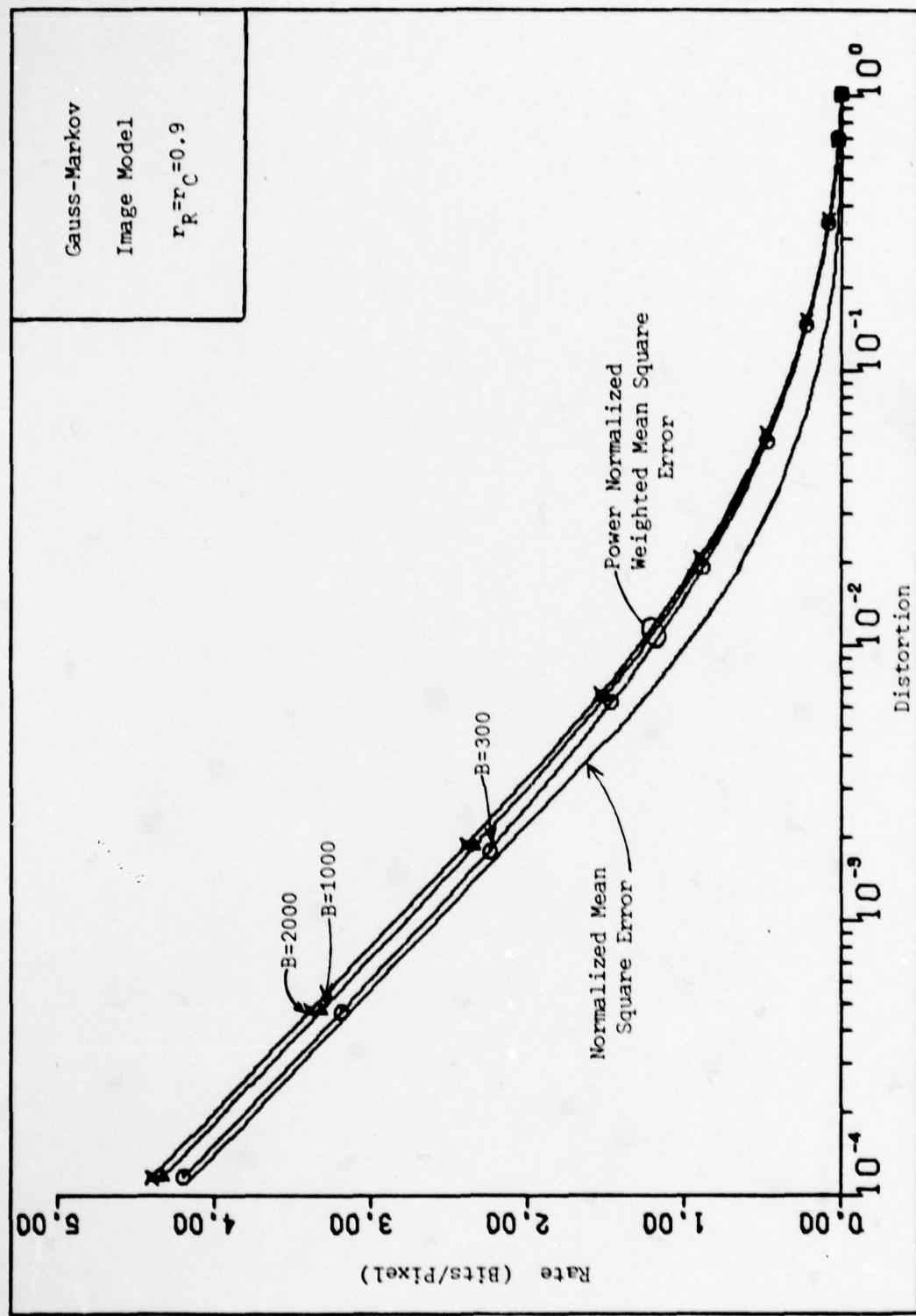


Figure 2.12 Dependence of $R(D)$ on Average Lumination (B) for Power Normalized WMSE

III. Performance of a Source Encoded Imagery Transmission System

Knowing the rate distortion function for a particular data source is useful in that it can provide a bench mark against which coding procedures for that source can be judged. Unfortunately, calculating the rate distortion function gives no information concerning the coding methods which should be applied to even get close to the $R(D)$ limit.

A considerable amount of research has been conducted in the area of image coding procedures, particularly in the realm of source encoding. Very little research has appeared on the use of channel encoding to reduce errors in the received image data stream and on the trade offs between source data bits and error correction coding bits. The purpose of this thesis is to explore these trade offs. In order to concentrate on this purpose no variations were allowed in the source encoding procedures. Instead, a thorough search of current literature was conducted to determine the "best" source encoding procedure for the purposes of the study.

The criteria used to select the source encoding procedure are:

1. The source encoder must be one of the best currently available image compression techniques in terms of performance. In this context performance means the ability to transmit an image using a limited number of data bits (as low as 1/2 bit/pixel) while maintaining low distortion levels. Since the distortion criteria used by most researchers is mean square error, this measure was used to compare the image encoding techniques.
2. The procedure must allow the trade off of individual data bits for error correction coding bits.
3. The technique should be one that is currently in practical use.

A description of the source encoding procedure selected and a detailed performance calculation for it are contained in this chapter. A block diagram of the image transmission system is shown in Figure 3.1.

The source encoding procedure which was determined to be the best in accordance with the criteria listed above is the two dimensional discrete cosine transform (DCT) followed by block quantization of the transform coefficients. The block quantizer simply quantizes each of the DCT coefficients to a preestablished number of bits. The data stream from the quantizer is fed into a channel encoder. For the purposes of this section the channel encoder/decoder and transmission link will be treated as a "black box" which accepts the quantizers output and returns it with errors to the receiver. The details of the channel encoder are discussed in Chapter IV. The cosine transform coefficients are reconstructed in the receiver and inverse transformed to provide the output image.

Discrete Cosine Transform

Why is it the "Best"? In the last several years, the source encoding procedure for imagery which has come to the forefront of technology is the cosine transform. There are several basic reasons for this. First, as a class the transform techniques (cosine, Fourier, Hadamard, Walsh and Karhunen-Loeve) have been shown to out perform all of the spatial coding techniques (run-length encoding, delta modulation, bit plane encoding, statistical coding) with the exception of differential pulse code modulation (DPCM) (Ref 23:662-680). At the same time, the transform techniques handle bit errors better because the inverse transformation process distributes the effect of the error over many pixels with any one pixel being only slightly changed. For most spatial techniques and especially

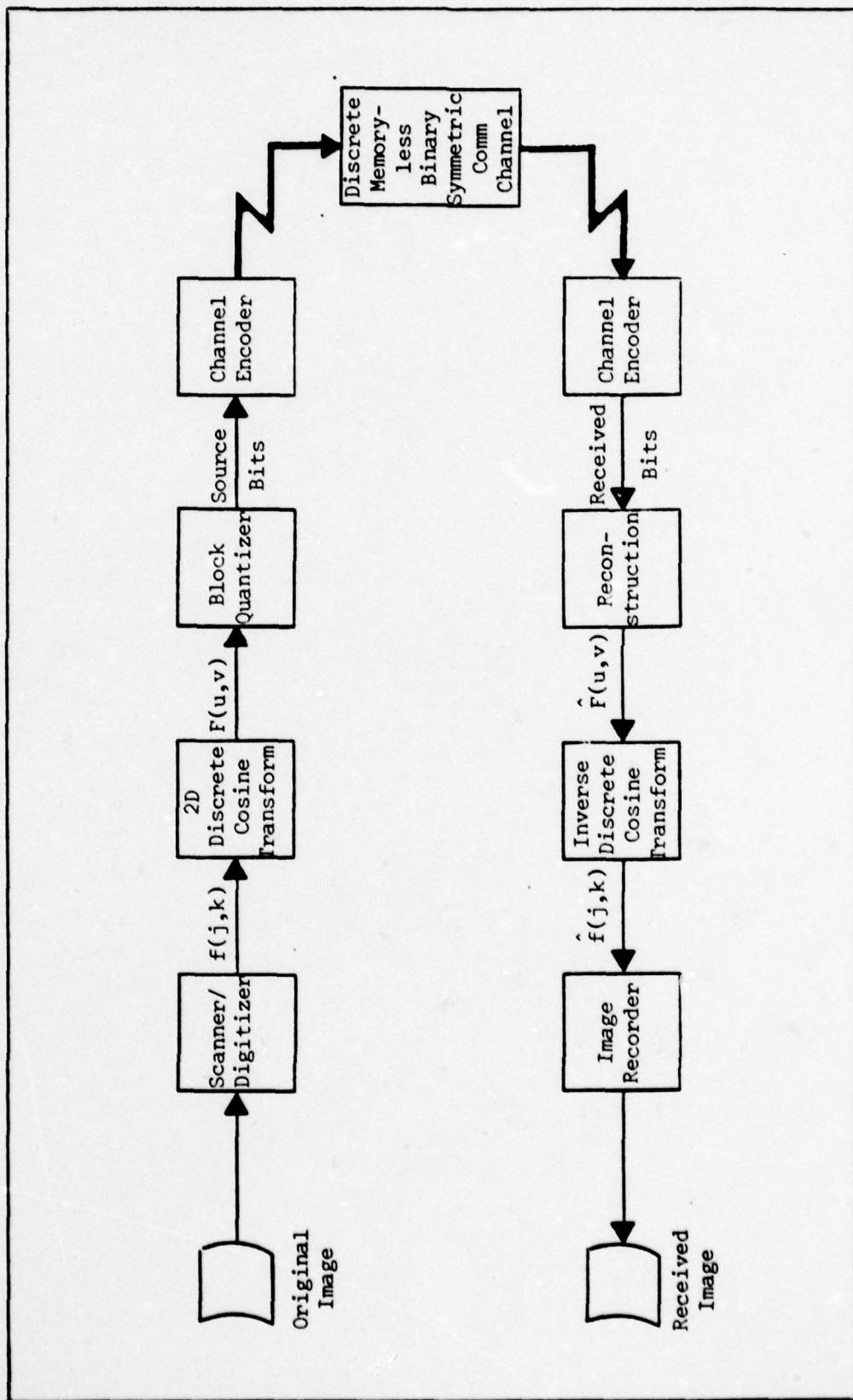


Figure 3.1 Source/Channel Encoded Digital Image Transmission System

in DPCM a single bit error can be propagated through many succeeding samples effectively wiping out all information until the next resynchronization pulses occur.

The basic justification for transform coding was offered by Huang and Schultheiss (Ref 20). They showed that the optimum coding procedure for a correlated Gaussian source consisted of transforming the data to a domain where the samples are uncorrelated and then optimally encoding each of the resulting samples using a memoryless coder. Of course, the transformation which accomplishes this is the Karhunen-Loeve (K-L) transform. Unfortunately, the K-L transform is very difficult to calculate since it depends entirely on the covariance structure of the source.

The cosine transform on the other hand is strictly a deterministic transform and can be implemented through very efficient "fast" methods. At the same time, it has been demonstrated that the cosine transform is virtually identical to the K-L transform in performance for many practical imagery applications (Ref 40:38). Table III.1 and Figure 3.2 show a comparison of the performance of several transform methods in the encoding of a one dimensional Gauss-Markov source as a function of the number of samples included in the transform. The cosine transform's performance is essentially that of the K-L transform. The figure also shows that there is little performance improvement for a transform size larger than 16. At the same time, the hardware size and complexity goes up sharply for larger block sizes. Therefore, a 16×16 block size was chosen for implementation in the simulated transmission system and for use in the theoretical performance calculation at the end of this chapter.

The final elements which have contributed to the rise of the cosine transform are the development of fast computational algorithms for it,

Table III.1

Performance of Various Transforms in Encoding a
1-D Gauss-Markov Source, $r=0.9$ (Ref 1:92)

Transform	M	(Mean Square Error)					
		2	4	8	16	32	64
Karhunen-Loeve	0.3730	0.2915	0.2533	0.2356	0.2268	0.2224	
Discrete Cosine	0.3730	0.2920	0.2546	0.2374	0.2282	0.2232	
Discrete Fourier	0.3730	0.2964	0.2706	0.2592	0.2441	0.2320	
Walsh-Hadamard	0.3730	0.2942	0.2649	0.2582	0.2582	0.2559	
Haar	0.3730	0.2942	0.2650	0.2589	0.2582	0.2581	

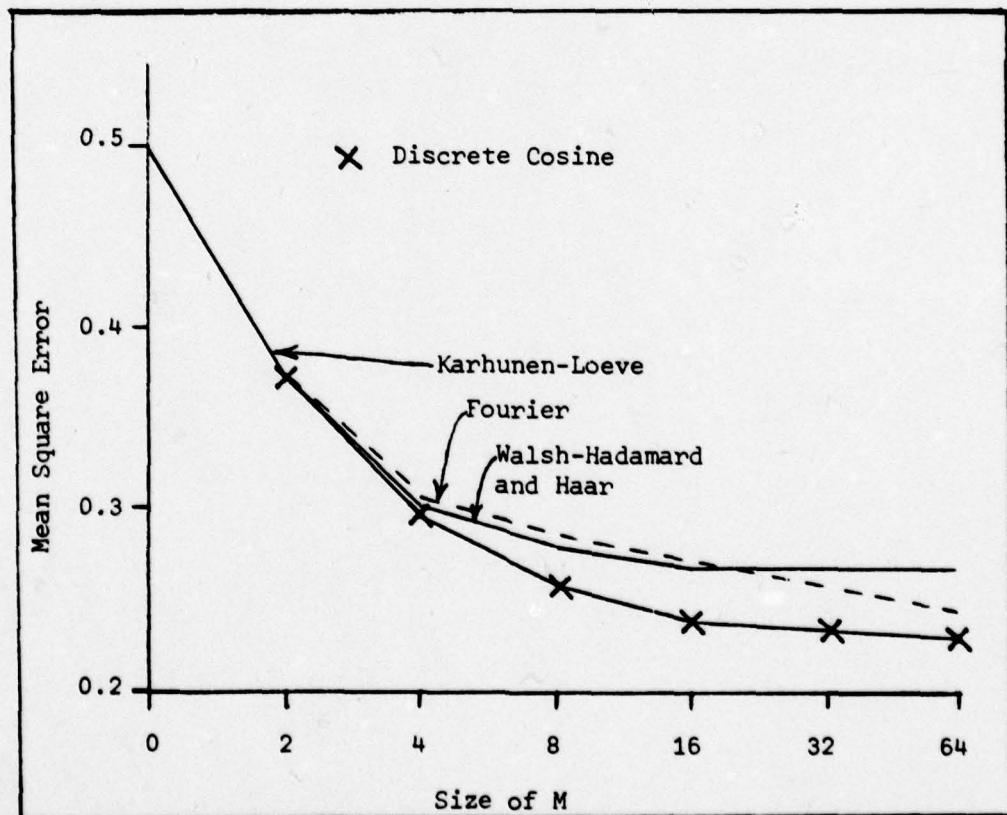


Figure 3.2 Performance of Various Transforms in Encoding a
1-D Gauss-Markov Source, $r=0.9$ (Ref 1:92)

and the ability to implement the transform in very compact hardware. The cosine transform was originally developed to solve the boundary discontinuity problems (known as the Gibbs phenomenon) associated with the Fourier transform. Doubling the number of points in the Fourier transform in a specific symmetric pattern eliminates the Gibbs Phenomenon (Ref 40:35-38). This double length Fourier transform has been designated the cosine transform. As a result of this historical development, the cosine transform has until recently been calculated using double length Fast Fourier Transform (FFT) techniques. In 1977 Chen, Smith and Fralick published a fast computational algorithm for the discrete cosine transform which they claim to be six times faster than the double length FFT (Ref 9). This method was implemented in the simulated image link and is described in detail in the next section.

Two additional benefits offered by the DCT are: (1) It operates only on real numbers as opposed to the Fourier transform which uses complex arithmetic. (2) The DCT can be implemented in very compact hardware using charge coupled devices or surface acoustic wave technology to perform the transform in analog form.

Computation of the Two Dimensional Cosine Transform. The two dimensional discrete cosine transform coefficients, $F(u,v)$, of an $M \times M$ array of image samples, $f(j,k)$, are defined as follows:

$$F(u,v) = \frac{4}{M^2} c(u)c(v) \sum_{j=0}^{M-1} \sum_{k=0}^{M-1} f(j,k) \cos\left[\frac{(2j+1)u\pi}{2M}\right] \cos\left[\frac{(2k+1)v\pi}{2M}\right]$$

$$(u,v) = 0, 1, 2, \dots M-1 \quad (3.1)$$

where

$$c(x) = \begin{cases} 1/\sqrt{2} & , x = 0 \\ 1 & , x = 1, 2, 3, \dots M-1 \end{cases} \quad (3.2)$$

The inverse transform is given by

$$f(j,k) = \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} c(u)c(v)F(u,v) \cos\left[\frac{(2j+1)u\pi}{2M}\right] \cos\left[\frac{(2k+1)v\pi}{2M}\right] \quad (j,k) = 0, 1, 2, \dots M-1 \quad (3.3)$$

Due to the separability of the DCT basis functions the two dimensional transform can be calculated by first transforming the data in one direction and then transforming the resultant one dimensional coefficients in the other direction as shown in the next equation.

$$F(u,v) = \frac{2}{M} c(u) \sum_{j=0}^{M-1} \cos\left[\frac{(2j+1)u\pi}{2M}\right] \left\{ \frac{2}{M} c(v) \sum_{k=0}^{M-1} f(j,k) \cos\left[\frac{(2k+1)v\pi}{2M}\right] \right\} \quad (3.4)$$

Each of the one dimensional DCT's of equation (3.4) can be calculated using the fast algorithm designed by Chen, Smith and Fralick (Ref 9). A signal flow graph representation of the 16 point algorithm appears in Figure 3.3. The algorithm consists of a series of "butterfly" operations in which pairs of data points are multiplied by appropriate constants and summed. The algorithm can be generalized to any size M such that M is a power of 2. The outputs of the flow graph must be multiplied by $2/M$ to give the normalized DCT coefficients defined by equation (3.1). To perform the inverse transform, the direction of the signal flow is reversed, the DCT coefficients, $F(u,v)$, are introduced at the output and the image

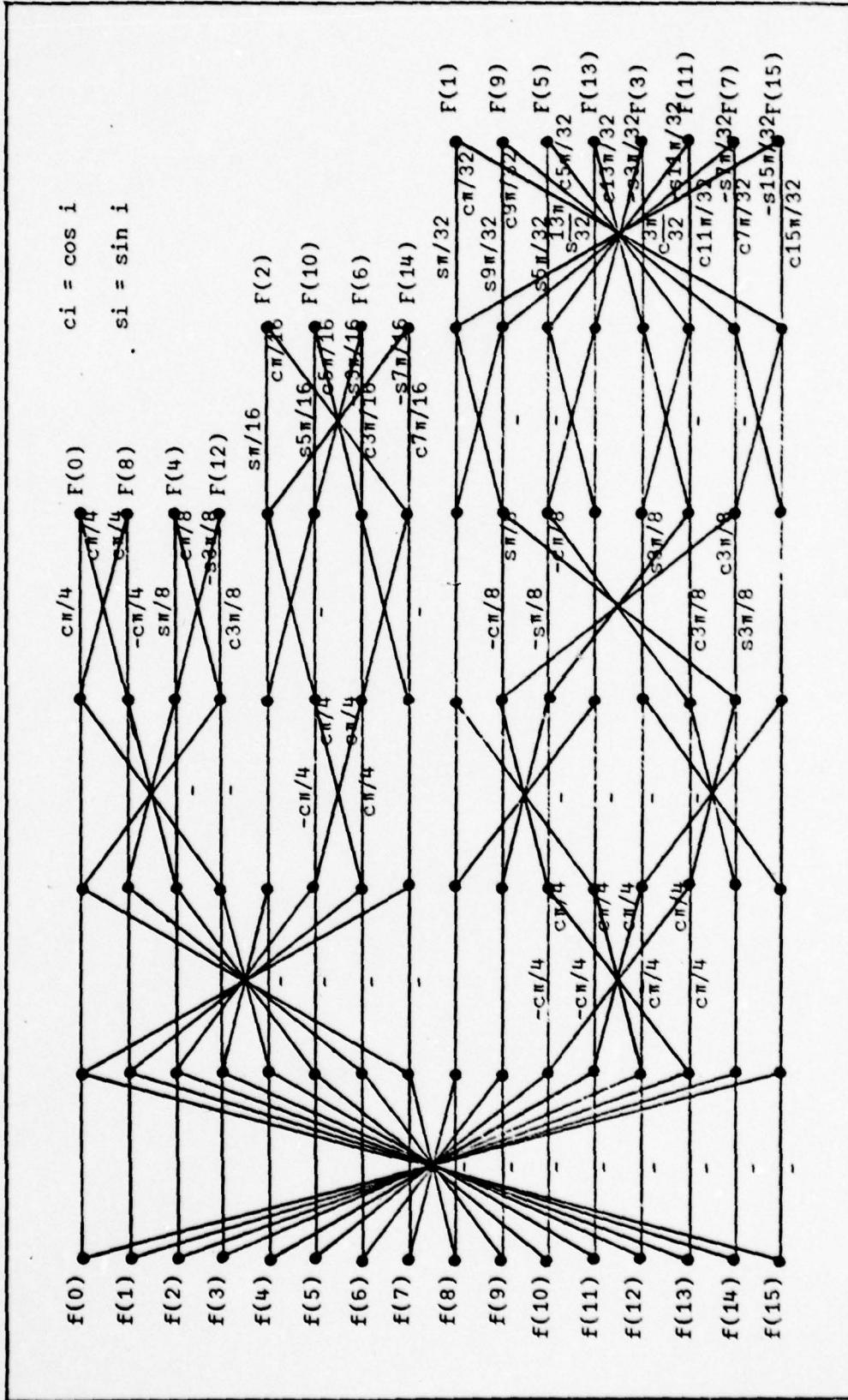


Figure 3.3 Fast Discrete Cosine Transform Signal Flow Graph, M=16 (Ref 9)

values, $f(j,k)$ are recovered at the input.

Statistical Properties of the Cosine Transform Coefficients. In calculating the performance of the source encoder and determining the optimum bit assignment for the block quantizer, the statistical properties of the DCT coefficients, $F(u,v)$, must be known.

Mean and Variance. The first parameters of interest are the mean and variance of $F(u,v)$. It is assumed that the image samples being fed to the DCT are drawn from a two dimensional Gauss-Markov process and that the image samples have been appropriately scaled so that they are 0 mean. The mean of the cosine transform sample under these assumptions is

$$\begin{aligned}
 E\{F(u,v)\} &= E\left\{\frac{4c(u)c(v)}{M^2} \sum_{j=0}^{M-1} \sum_{k=0}^{M-1} f(j,k) \cos\left[\frac{(2j+1)u\pi}{2M}\right] \cos\left[\frac{(2k+1)v\pi}{2M}\right]\right\} \\
 &= \frac{4c(u)c(v)}{M^2} \sum_{j=0}^{M-1} \sum_{k=0}^{M-1} E\{f(j,k)\} \cos\left[\frac{(2j+1)u\pi}{2M}\right] \cos\left[\frac{(2k+1)v\pi}{2M}\right] \\
 &= 0 \text{ for all } (u,v) \tag{3.5}
 \end{aligned}$$

Since the DCT coefficients are zero mean, their variances are given by

$$\begin{aligned}
 \sigma^2(u,v) &= E\{F^2(u,v)\} \\
 &= \frac{16c^2(u)c^2(v)}{M^4} \sum_{j=0}^{M-1} \sum_{k=0}^{M-1} \sum_{l=0}^{M-1} \sum_{m=0}^{M-1} E\{f(j,k)f(l,m)\} \cdot \\
 &\quad \cos\left[\frac{(2j+1)u\pi}{2M}\right] \cos\left[\frac{(2k+1)v\pi}{2M}\right] \cos\left[\frac{(2l+1)u\pi}{2M}\right] \cos\left[\frac{(2m+1)v\pi}{2M}\right] \tag{3.6}
 \end{aligned}$$

For the Gauss-Markov image model

$$E\{f(j,k)f(l,m)\} = \sigma_f^2 r_R^{|j-l|} r_C^{|k-m|} \quad (3.7)$$

where σ_f^2 is the common variance of the image samples and r_R and r_C are the row and column correlation coefficients, respectively. Substituting (3.7) into (3.6) gives

$$\sigma^2(u,v) = \frac{16\sigma_f^2 c^2(u)c^2(v)}{M^4} \sum_{j=0}^{M-1} \sum_{l=0}^{M-1} \sum_{m=0}^{M-1} \sum_{n=0}^{M-1} r_R^{|j-l|} r_C^{|k-m|} \quad (3.8)$$

$$\cos\left[\frac{(2j+1)u\pi}{2M}\right] \cos\left[\frac{(2k+1)v\pi}{2M}\right] \cos\left[\frac{(2l+1)u\pi}{2M}\right] \cos\left[\frac{(2m+1)v\pi}{2M}\right]$$

$$(u,v) = 0, 1, 2, 3, \dots M-1$$

The array of $\sigma^2(u,v)$ values for $M=8$, $r_R = r_C = 0.9$ and $\sigma_f^2 = 1$ was computed using equation (3.8) and is shown in Table III.2. It can be seen that most of the energy of the image data is concentrated into the low frequency coefficients of the DCT. This is the key to using the DCT to reduce the required band-width for transmission of the image data. Many of the high order DCT coefficients can simply be thrown away at the transmitter and replaced with zeros at the receiver with little effect on image quality but with a significant reduction of transmitted bits.

It should be pointed out that the DCT coefficients of the Gauss-Markov transformed data are not completely uncorrelated. Calculation of $E\{F(u,v)F(w,x)\}$ reveals that a few of the low frequency off diagonal terms of the correlation matrix are non-zero. However, these terms have a magnitude which is less than 10% of the variance of $F(u,v)$. Also as M increases or as r_R and r_C approach 1, the non-zero off diagonal terms

Table III.2
 Variances of DCT Coefficients $\{\sigma^2(u, v)\}$ of 2-D Gauss-Markov Image Model
 $(r_R = r_C = 0.9; M=8)$

<u>u</u> \ <u>v</u>	0	1	2	3	4	5	6	7
0	2.3913	.3889	.1338	.0641	.0404	.0293	.0238	.0211
1	.3889	.0632	.0218	.0104	.0066	.0048	.0039	.0034
2	.1338	.0218	.0075	.0036	.0023	.0016	.0013	.0012
3	.0641	.0104	.0036	.0017	.0011	.0008	.0006	.0006
4	.0404	.0066	.0023	.0011	.0007	.0005	.0004	.0004
5	.0293	.0048	.0016	.0008	.0005	.0004	.0003	.0003
6	.0238	.0039	.0013	.0006	.0004	.0003	.0002	.0002
7	.0211	.0034	.0012	.0006	.0004	.0003	.0002	.0002

get much smaller. Therefore the DCT does a very adequate job reducing the correlation in the image data. It will be assumed from here on that the DCT coefficients are uncorrelated.

Probability Density for the DCT Coefficients. The image samples, $f(j,k)$, are assumed to be drawn from a Gaussian process, therefore they will be jointly Gaussian. Since the DCT coefficients are linear combinations of the image samples, they will also be jointly Gaussian (Ref 29:222). A fundamental result of probability theory states that if random variables are jointly Gaussian and uncorrelated, they are independent (Ref 29:221). The DCT coefficients are assumed to be both Gaussian and uncorrelated so they will be assumed to be independent. Therefore, the probability density $P_{u,v}(x)$ for the (u,v) coefficient is given by

$$P_{u,v}(x) = \frac{1}{\sqrt{2\pi} \sigma(u,v)} \exp\left\{-\frac{x^2}{2\sigma^2(u,v)}\right\} \quad (3.9)$$

where $\sigma^2(u,v)$ is defined by equation (3.8)

These statistical characteristics are critical to the design of the next stage in the image transmission system, the block quantizer.

Block Quantization of the DCT Coefficients

The optimal block quantizer would be designed to minimize the total reconstructed error in the received image. Unfortunately, it is difficult to relate the quantization scheme to the total error since this would require exact knowledge of the communications error patterns to be encountered. Rather than design the quantizer to minimize the total error, the approach taken in this thesis is to construct the block quantizer in a manner which will minimize the quantization error. At first glance, this appears to be a decidedly sub-optimum approach. However, the application of effective

channel coding methods can make the quantizer the only significant source of error in the system.

The Shannon channel coding theorem states that as long as the transmission rate does not exceed the channel's capacity, error correcting codes and appropriate decoding rules exist which allow the probability of decoder error to be as small as desired (Ref 25:81). In the typical military communications system the data transmission rates are far below the theoretical channel capacity. Therefore, Shannon's channel coding theorem guarantees that the distortion in the reconstructed image due to channel noise can be reduced to insignificant levels. The only remaining source of distortion in the image transmission system will be the quantization error.

Under these circumstances, the problem of designing the block quantizer can be reduced to deciding how many bits or equivalently how many quantization levels should be assigned to each of the DCT coefficients in order to minimize the total quantization error. To complete the quantizer design, the distribution of the quantization levels over the possible range of the individual DCT terms must be determined.

Bit Assignment for the Block Quantizer. The bit assignment problem can be expressed in two equivalent ways: (1) Given a fixed number of bits (T) to be assigned to the M^2 DCT samples (i.e. if an average transmission rate of T/M^2 bits/pixel is desired) what is the optimum procedure for dividing these T bits among the M^2 samples so as to achieve the lowest distortion? (2) Given a desired distortion level (D) what is the minimum transmission rate (R) in bits per pixel which will guarantee that the average quantization error is less than or equal to D ? Expressing the problem in this second manner immediately leads to the conclusion that the

desired rate (R) is given by the rate distortion function for the cosine transform coefficients, $R_C(D)$. Further, it will be shown that the particular structure of $R_C(D)$ allows it to be used to define the optimum bit assignments in the block quantizer.

The assumed statistical properties of the DCT coefficients are critical to the calculation of $R_C(D)$. Berger (Ref 5:57) has shown that for a set of M^2 statistically independent discrete memoryless sources with rate distortion functions $R_{u,v}(D_{u,v})$ the total rate distortion function $R_C(D)$ is given by

$$R_C(D) = \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} R_{u,v}(D_{u,v}) \quad (3.10)$$

if

$$D = \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} D_{u,v} \quad (3.11)$$

Since the DCT coefficients are approximately independent, equation (3.10) will be used to define the rate distortion function for them.

To completely specify $R_C(D)$, a definition is needed for $D_{u,v}$, the distortion measure in the frequency domain that corresponds to the distortion measure defined by (2.8) in the spatial domain. Equation (2.8) is repeated here for convenience. The expectation operator has been omitted to simplify the following equations.

$$D = \frac{1}{M^2} \sum_{j=0}^{M-1} \sum_{k=0}^{M-1} \left[a(j,k) * \left(f(j,k) - \hat{f}(j,k) \right) \right]^2 \quad (3.12)$$

where

$a(j,k)$ = spatial weighting function

$f(j,k)$ = original image sample values

$\hat{f}(j,k)$ = reconstructed image samples

Figure 3.4 gives a physical interpretation of this distortion measure.

If $a(j,k) * f(j,k)$ is defined to be $g(j,k)$ and $a(j,k) * \hat{f}(j,k)$ defined to be $\hat{g}(j,k)$, the distortion measure becomes

$$D = \frac{1}{M^2} \sum_{j=0}^{M-1} \sum_{k=0}^{M-1} [g(j,k) - \hat{g}(j,k)]^2 \quad (3.13)$$

which is just the mean square error in what will be called the "visual domain". The distortion measure simply finds the squared difference between the original image samples and the reconstructed image samples after they have been filtered by the human visual system.

It is important to remember that the form of $a(j,k)$ has not been directly determined but instead has been inferred from its frequency domain version $A(u,v)$. Furthermore the characteristics of $A(u,v)$ have been measured by presenting single frequency sinusoidal displays to human observers. The effect of the cosine transform is to break the visual domain data into its constituent single frequency components. Therefore, it is conjectured that measuring the distortion in the spatial domain as shown in Figure 3.4 is equivalent to measuring it in the frequency domain as shown in Figure 3.5. That is

$$D = \frac{1}{M^2} \sum_{j=0}^{M-1} \sum_{k=0}^{M-1} [g(j,k) - \hat{g}(j,k)]^2 \quad (3.14)$$

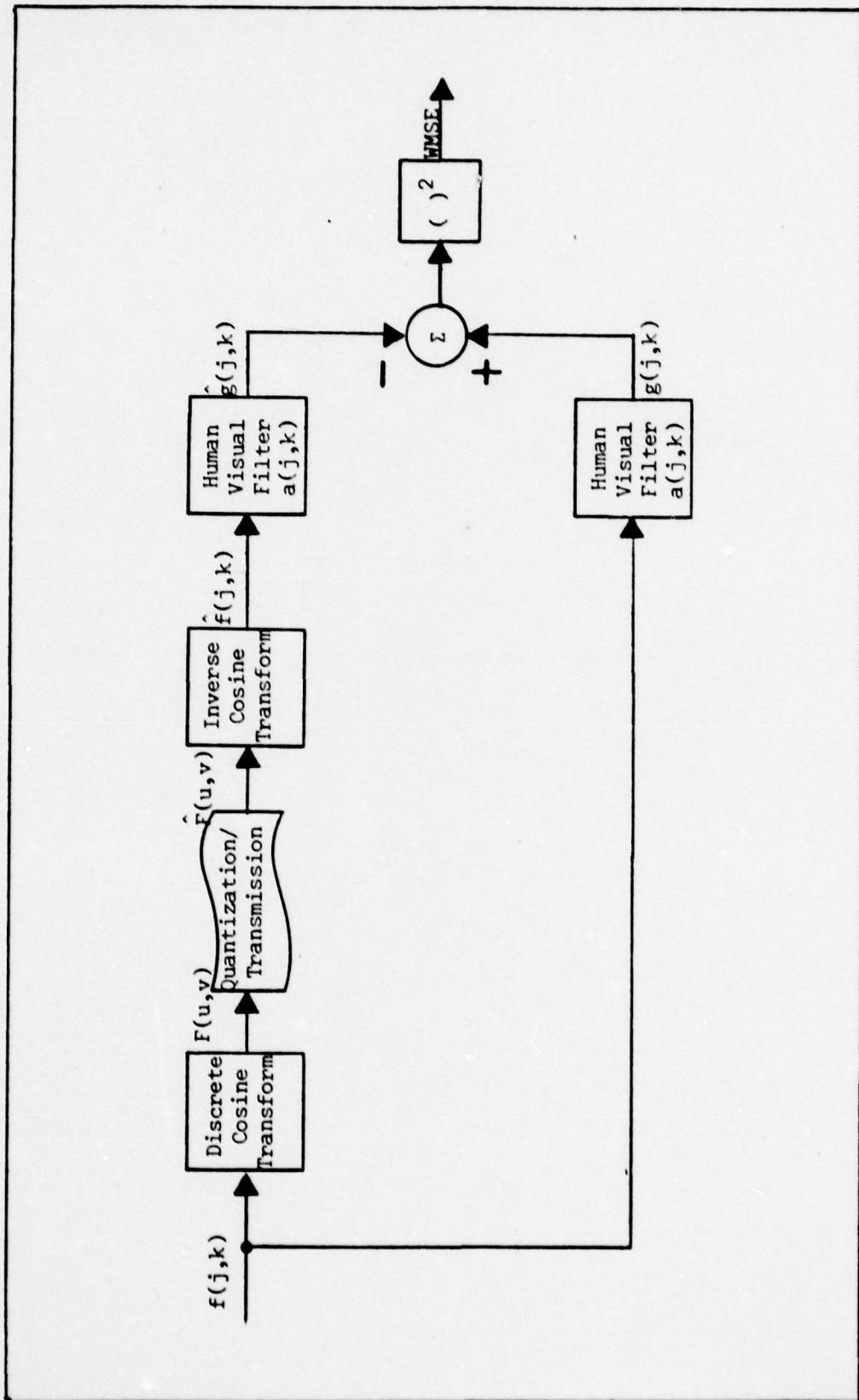


Figure 3.4 Weighted Mean Square Error in the Spatial Domain

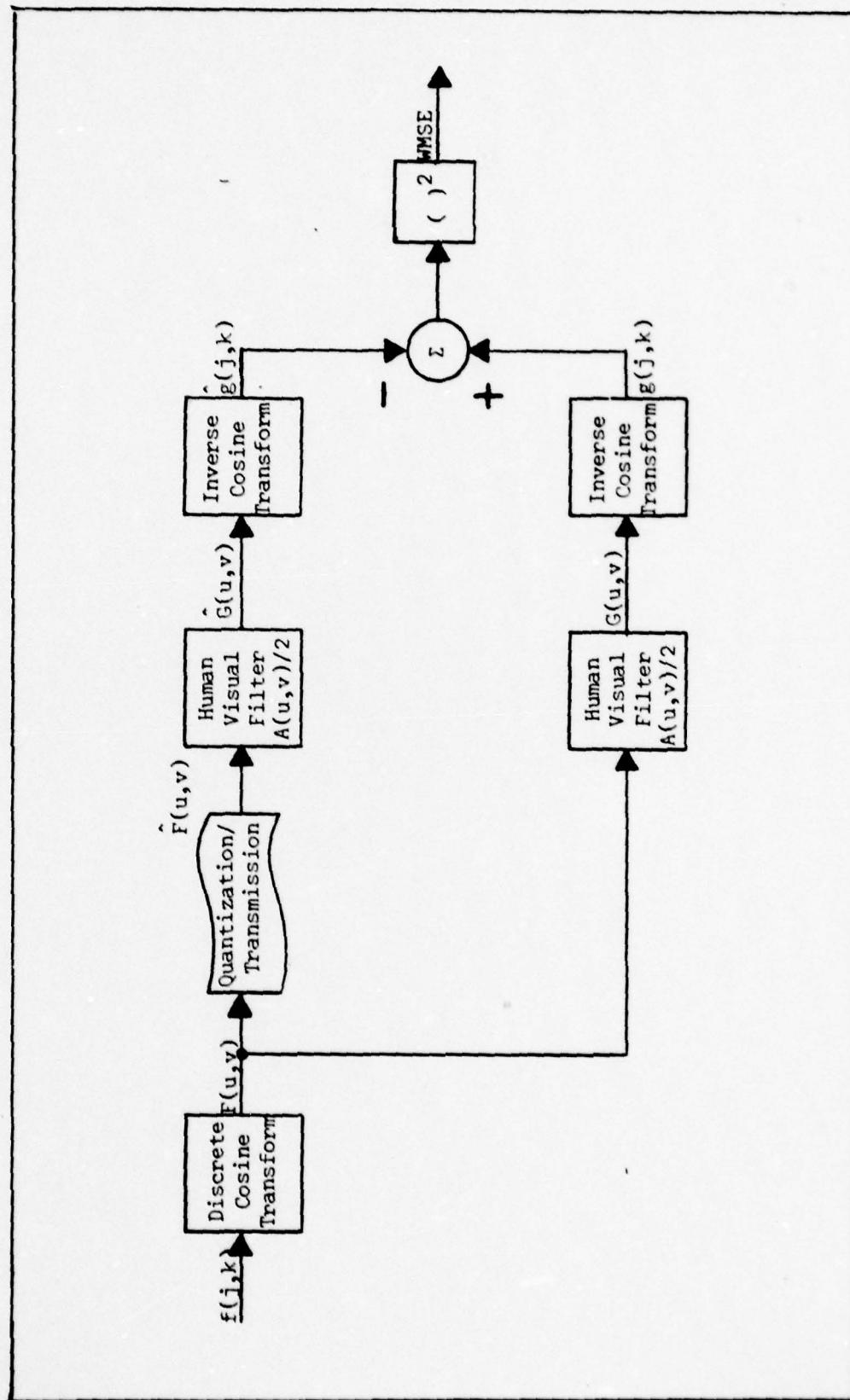


Figure 3.5 Weighted Mean Square Error in the Cosine Transform Domain

$$\begin{aligned}
 &= \frac{1}{M^2} \sum_{j=0}^{M-1} \sum_{k=0}^{M-1} \left[a(j,k) * (f(j,k) - \hat{f}(j,k)) \right]^2 \\
 &= K \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} A^2(u,v) \left[F(u,v) - \hat{F}(u,v) \right]^2 \\
 &= K \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} \left[G(u,v) - \hat{G}(u,v) \right]^2
 \end{aligned} \tag{3.14}$$

where

$$\begin{aligned}
 K &= \text{proportionality constant} \\
 G(u,v) &= A(u,v) F(u,v) \\
 \hat{G}(u,v) &= A(u,v) \hat{F}(u,v)
 \end{aligned}$$

Equation (3.14) can be shown to be true analytically for the Fourier transform since convolution in the spatial domain is equivalent to multiplication in the frequency domain. However, the frequency terms $F(u,v)$ and $\hat{F}(u,v)$ in equation (3.14) are not Fourier frequency terms. They are cosine transform samples and no convolution type theorem for the cosine domain appears to exist. Therefore at this point it will be assumed that equation (3.14) holds. It will be shown later that calculation of the cosine transform rate distortion function appears to confirm the validity of equation (3.14).

In order to determine K , the proportionality constant in equation (3.14), the definitions of $g(j,k)$ and $\hat{g}(j,k)$ in terms of their cosine transforms $G(u,v)$ and $\hat{G}(u,v)$ are substituted into (3.14)

$$D = \frac{1}{M^2} \sum_{j=0}^{M-1} \sum_{k=0}^{M-1} \left\{ \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} c(u)c(v) [G(u,v) - \hat{G}(u,v)] \cdot \right. \quad (3.15)$$

$$\left. \cos\left[\frac{(2j+1)u\pi}{2M}\right] \cos\left[\frac{(2k+1)v\pi}{2M}\right] \right\}^2$$

Expanding the squared term into a double summation gives

$$D = \frac{1}{M^2} \sum_{j=0}^{M-1} \sum_{k=0}^{M-1} \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} \sum_{w=0}^{M-1} \sum_{x=0}^{M-1} c(u)c(v)c(w)c(x) \cdot$$

$$[G(u,v) - \hat{G}(u,v)][G(w,x) - \hat{G}(w,x)] \cos\left[\frac{(2j+1)u\pi}{2M}\right] \cos\left[\frac{(2k+1)v\pi}{2M}\right] \cdot \quad (3.16)$$

$$\cos\left[\frac{(2j+1)w\pi}{2M}\right] \cos\left[\frac{(2k+1)x\pi}{2M}\right]$$

Rearranging terms and reversing the order of summation gives

$$D = \frac{1}{M^2} \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} \sum_{w=0}^{M-1} \sum_{x=0}^{M-1} [G(u,v) - \hat{G}(u,v)][\hat{G}(w,x) - G(w,x)] \cdot$$

$$c(u)c(w) \sum_{j=0}^{M-1} \cos\left[\frac{(2j+1)u\pi}{2M}\right] \cos\left[\frac{(2j+1)w\pi}{2M}\right] \cdot \quad (3.17)$$

$$c(v)c(x) \sum_{k=0}^{M-1} \cos\left[\frac{(2k+1)v\pi}{2M}\right] \cos\left[\frac{(2k+1)x\pi}{2M}\right]$$

The last two terms of equation (3.17) can be shown to evaluate to $M/2$ for $u=w$ (or $v=x$) and 0 for $u \neq w$ (or $v \neq x$). Therefore (3.17) reduces to

$$D = \frac{1}{M^2} \left\{ \frac{M^2}{4} \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} [G(u,v) - \hat{G}(u,v)]^2 \right\} \quad (3.18)$$

Therefore the proportionality constant k equals $1/4$ and the distortion measure in the frequency domain is

$$D = E \left\{ \frac{1}{4} \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} A^2(u,v) [F(u,v) - \hat{F}(u,v)]^2 \right\} \quad (3.19)$$

Returning to the rate distortion function, $R_C(D)$, given in equation (3.10), if $D_{u,v}$ is defined to be

$$D_{u,v} = E \left\{ \frac{A^2(u,v)}{4} [F(u,v) - \hat{F}(u,v)]^2 \right\} \quad (3.20)$$

then the criterion of (3.11) is met and the total rate distortion function for the DCT coefficients is approximately the sum of the rate distortion functions for the individual terms, $R_{u,v}(D_{u,v})$.

$R_{u,v}(D_{u,v})$ is the rate distortion function of a memoryless Gaussian source, $F(u,v)$, with mean 0 and variance $\sigma^2(u,v)$ under a weighted mean square error distortion criteria, where the weight is $A^2(u,v)/4$. If equation (3.20) is rearranged slightly to give

$$D_{u,v} = E \left\{ \left[\frac{A(u,v)}{2} F(u,v) - \frac{A(u,v)}{2} \hat{F}(u,v) \right]^2 \right\} \quad (3.21)$$

it can be seen that $R_{u,v}(D_{u,v})$ is simply the rate distortion function of a memoryless Gaussian source with variance $A^2(u,v)\sigma^2(u,v)/4$ under an unweighted mean square error criteria. $R(D)$ for this case is given by equation (2.14)

$$R_{u,v}(D_{u,v}) = \begin{cases} \frac{1}{2} \log \frac{A^2(u,v)\sigma^2(u,v)}{4D_{u,v}}, & 0 \leq D_{u,v} \leq \frac{A^2(u,v)\sigma^2(u,v)}{4} \\ 0, & D_{u,v} \geq \frac{A^2(u,v)\sigma^2(u,v)}{4} \end{cases} \quad (3.22)$$

Substituting (3.22) into (3.10) reveals that the total rate distortion function for the cosine transform source $R_C(D)$ is given parametrically by

$$R_C[D(Q)] = \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} \max \left\{ 0, \frac{1}{2} \log \left[\frac{A^2(u,v)\sigma^2(u,v)}{4Q} \right] \right\} \quad (3.23)$$

$$D(Q) = \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} \min \left\{ Q, \frac{A^2(u,v)\sigma^2(u,v)}{4} \right\}$$

This rate distortion function has been calculated numerically and a comparison of it to the rate distortion function for the image source is shown in Figure 3.6. Both curves were calculated using the power-normalized weighting function $A_p(w_1, w_2)$ defined by equation (2.24). It can be seen that the DCT rate distortion function very closely approximates the image rate distortion function. The small difference between them is a manifestation of the correlation which remains between the DCT coefficients. If this correlation could be exploited the two curves would be still closer. Figure 3.6 also gives support to the assumption that the distortion criterion in the frequency domain given by equation (3.19) is equivalent to the spatial domain distortion measure given by (3.12). The cosine rate distortion function was calculated using (3.19) while the image rate distortion function used (3.12).

The reason for finding $R_C(D)$ is that it can be used to define the

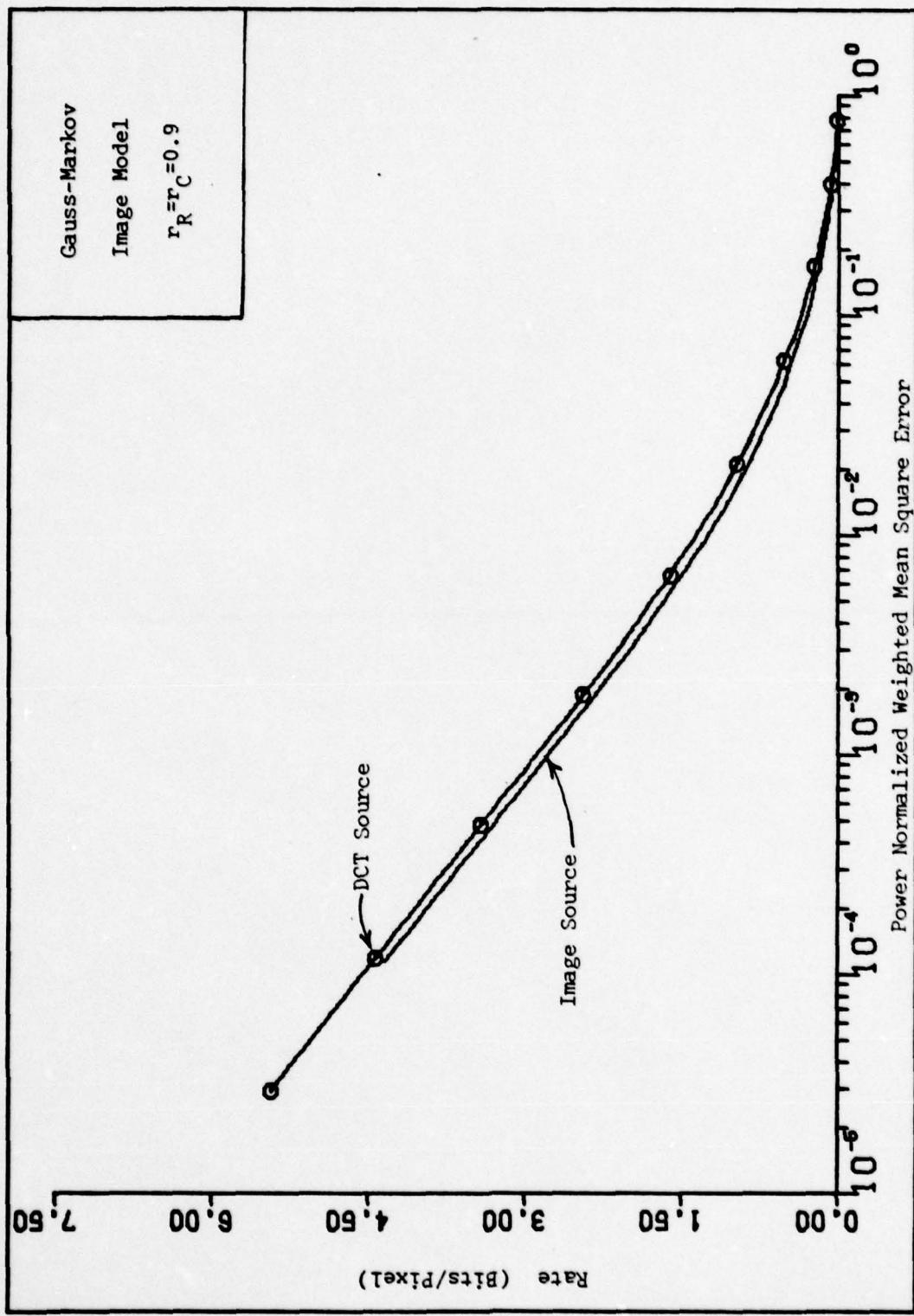


Figure 3.6 Comparison of Rate Distortion Functions for Image and DCT Sources

number of bits to assign to each DCT coefficient in order to minimize the quantization error. Since the total rate distortion function is merely the sum of the individual coefficient rate distortion functions, the optimum number of bits, $n(u,v)$, to be assigned to the (u,v) coefficient is given by

$$n(u,v) = \begin{cases} R_{u,v}(Q) = \frac{1}{2} \log_2 \frac{A^2(u,v)\sigma^2(u,v)}{4Q}, & \frac{A^2(u,v)\sigma^2(u,v)}{4} > Q \\ 0, & \frac{A^2(u,v)\sigma^2(u,v)}{4} < Q \end{cases} \quad (3.24)$$

This assignment rule must be modified, however, to accommodate the fact that $R_{u,v}(Q)$ is a real number while $n(u,v)$ must be an integer for practical application. Various authors have suggested algorithms for assigning a fixed number of bits (T) to M^2 coefficients (Ref 32:150;34:79;44:650). All procedures basically rely on the following procedure:

- 1.) Compute the bit assignment $n(u,v)$ from (3.24)
- 2.) Round off each $n(u,v)$ to the nearest integer
- 3.) If $\sum_{u=0}^{M-1} \sum_{v=0}^{M-1} n(u,v) \neq T$, "adjust" some of the $n(u,v)$ until the total number of bits does equal T .

Variation in the algorithms centers on what rules should be used to "adjust" the number of bits in step 3. In the simulation discussed in Chapter V, the adjustments were made to produce the lowest weighted mean square error based on the theoretical performance calculation completed in this chapter.

Several example bit assignment matrices are shown in Table III.3 for both the weighted MSE and the unweighted MSE distortion criteria. The effect of the frequency weighting can be seen quite clearly in the $R = 2.04$ case. The MSE criterion would assign 8 bits to the $(0,0)$ coefficient.

Table III.3 Bit Assignment Matrices for the DCT Coefficients

The WMSE measure assigns only 5. The weighting function shifts the emphasis away from the very low frequency components and spreads the bits more evenly among the mid-frequency components. For both criteria the high frequency components are assigned very few bits.

Quantization Strategy for Individual DCT Terms. Once the number of bits, $n(u,v)$, to be assigned to each DCT coefficient has been determined, the next step that must be taken is to determine how to distribute the $L(u,v)$ levels $\left[L(u,v) = 2^{n(u,v)} \right]$ over the range of the coefficient. Max has given the optimal method for quantizing a Gaussian random variable with zero mean and unit variance in his 1960 article "Quantizing for Minimum Distortion" (Ref 24). The optimality condition for Max was minimum mean square error. The Max results can be used in the quantization of the DCT coefficients if the source being quantized is defined to be the weighted DCT coefficient $A(u,v) F(u,v)$ divided by its standard deviation $A(u,v) \sigma(u,v)$.

$$F_n(u,v) = \frac{A(u,v) F(u,v)}{A(u,v) \sigma(u,v)} = \frac{F(u,v)}{\sigma(u,v)} \quad (3.25)$$

Since $F(u,v)$ is zero mean $F_n(u,v)$ will be also and the variance of $F_n(u,v)$ will be unity.

In his article, Max also demonstrated that the optimal quantization method gave only a small performance gain over the sub-optimal uniform quantization method. The structure of the uniform quantizer is shown in Figure 3.7. The width of the quantizer bins (d) is calculated to minimize the total mean square error. It is important in this application that the uniform quantizer's performance is relatively close to that of the optimal quantizer. In the optimal quantization procedure the endpoints of the

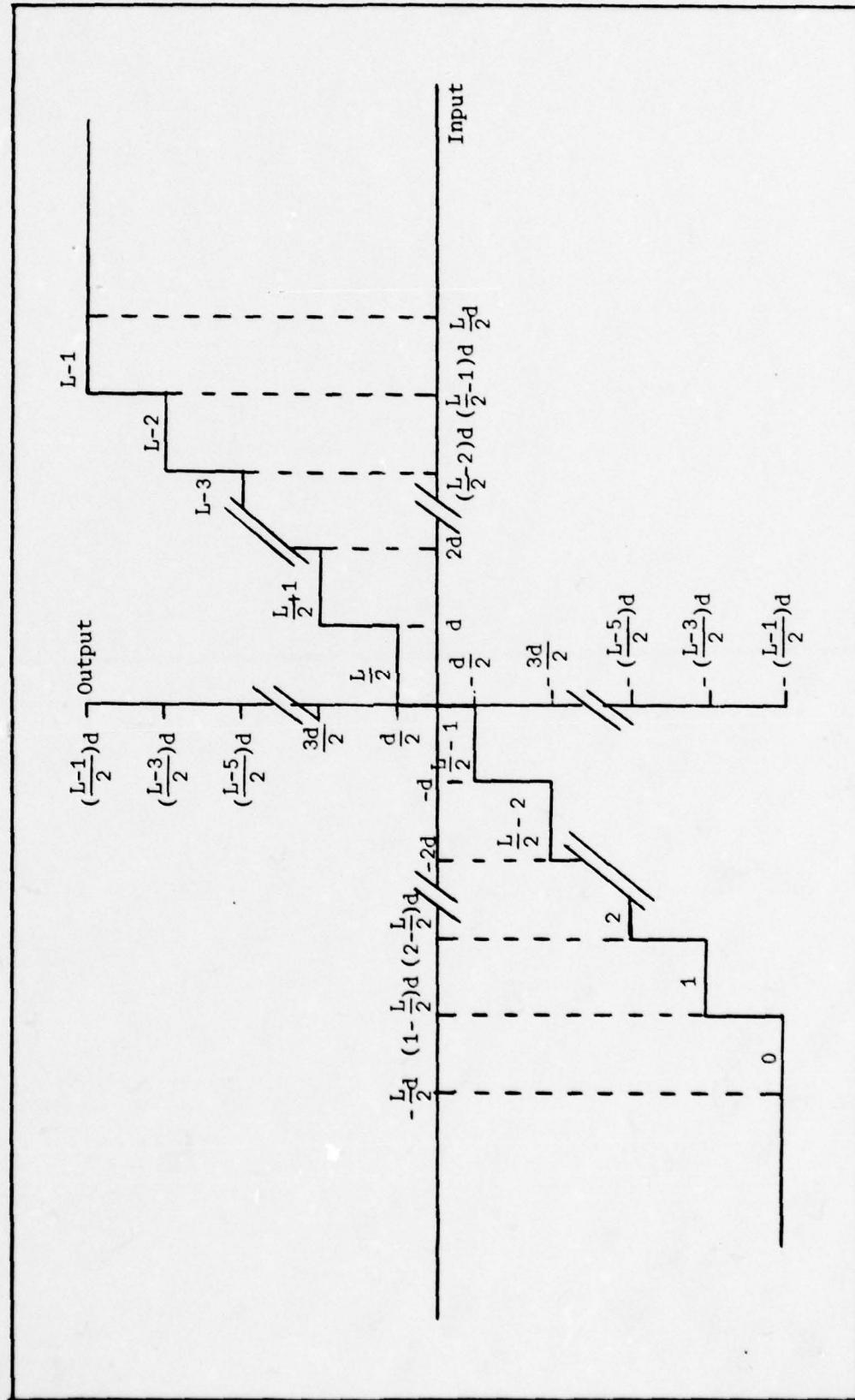


Figure 3.7 Structure of the Uniform Quantizer with $L=2^n$ levels ($d=\text{optimal level spacing}$)

quantization intervals are not uniformly spaced and must be precomputed and stored in a table. Also the output value for each quantizer level is not the mid-point of the interval but instead is the center of probability mass for the interval so it, too, must be precomputed and stored. The bit assignment procedure described in the last section can sometimes result in as many as 16 bits being assigned to the low frequency DCT coefficients. This implies that for the optimal quantizer more than 260,000 constants must be stored. For the uniform quantizer only the value of d is required to define the quantization interval endpoints and output levels. Therefore only 16 values of d must be stored, one for each possible value of $n(u,v)$. For this reason, the uniform quantizer was used in the theoretical performance calculation and the image simulation.

Max gave values for d and the resulting MSE only up to 36 output levels (Ref 24). As mentioned above the bit assignment process for the DCT coefficients can require as many as 16 bits or 65,536 levels. Therefore it was necessary to extend the Max results to include all 2^n number of output levels, $n = 1, 2, 3, \dots, 16$. The defining equation which must be minimized to find the optimum value of d is (ref 24:9)

$$D = 2 \sum_{i=1}^{L/2-1} \int_{(i-1)d}^{id} \left[x - \left(\frac{2i-1}{2} \right) d \right]^2 p(x) dx + 2 \int_{(L/2-1)d}^{\infty} \left[x - \left(\frac{L-1}{2} \right) d \right]^2 p(x) dx \quad (3.26a)$$

where L is the number of quantizer levels and $p(x)$ is the $N(0,1)$ probability density. To minimize (3.26a) with respect to d , the partial derivative of D with respect to d is taken and the resulting equation set equal to zero.

$$\frac{\partial D}{\partial d} = - \sum_{i=1}^{L/2-1} (2i-1) \int_{(i-1)d}^{id} \left[x - \left(\frac{2i-1}{2} \right) d \right] p(x) dx$$

(3.26b)

$$- (L-1) \int_{(L/2-1)d}^{\infty} \left[x - \left(\frac{L-1}{2} \right) d \right] p(x) dx = 0$$

Using numerical search techniques the root (d) of equation (3.26b) was found for each $L = 2^n$, $n = 1, 2, \dots, 16$. These d values were then substituted into equation (3.26a) to find the resultant value for the mean square error. Table III.4 gives the calculated d and MSE values.

Representation of the Quantizer Output Level. The final issue concerning the block quantizer that must be addressed is what binary code should be used to represent the output of the quantizer. Some limited research on this subject was conducted but little information was found that would cause one code to be preferred to another. Therefore for ease of implementation and compatibility with standard computer formats the natural binary code in which an integer k is represented by

$$k = \sum_{i=0}^{n-1} k_i 2^i \quad (3.27)$$

where

$$k_i = 0, 1$$

was selected for use in the following performance calculation and the image transmission simulation.

Theoretical Performance of the Source Encoded Imagery Transmission System

The objective of this section is to develop an expression for the total expected distortion in the reconstructed image (D_T). The distortion

Table III.4
Optimum Level Spacing and Resultant Mean
Square Error for the Uniform Quantizer

<u>Number of Bits</u>	<u>Number of Levels</u>	<u>d</u>	<u>MSE</u>
1	2	1.59576912	.36338023E+00
2	4	.99568669	.11884605E+00
3	8	.58601944	.37439661E-01
4	16	.33520061	.11542885E-01
5	32	.18813879	.34952114E-02
6	64	.10406301	.10400449E-02
7	128	.05686767	.30433305E-03
8	256	.03076239	.87686339E-04
9	512	.01649895	.24919336E-04
10	1024	.00878546	.69972886E-05
11	2048	.00464983	.19447093E-05
12	4096	.00244839	.53582554E-06
13	8192	.00128358	.14667211E-06
14	16384	.00067037	.40047439E-07
15	32768	.00034892	.11043099E-07
16	65536	.00018097	.32263614E-08

measure will be the weighted mean square error criterion defined by (2.8). As shown in the previous section this distortion criterion can be written in the discrete cosine transform domain as (3.19). Since the weighting function is deterministic the expectation operator can be taken inside the summation of (3.19) to give

$$D_T = \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} \frac{A^2(u,v)}{4} E \left\{ \left[F(u,v) - \hat{F}(u,v) \right]^2 \right\} \quad (3.28)$$

Therefore the task of determining D_T reduces to calculating the mean square error between the original and reconstructed values for the individual cosine transform coefficients.

There are two sources of variation between $F(u,v)$ and $\hat{F}(u,v)$. First is the quantization process. Second is the error induced by noise in the communications channel. These effects will be modeled as additive. That is

$$\hat{F}(u,v) = F(u,v) + Q(u,v) + N(u,v) \quad (3.29)$$

$Q(u,v)$ = quantization error

$N(u,v)$ = error contribution due to channel noise

Further it will be assumed that the two error terms are independent. Therefore

$$\begin{aligned} E \left\{ \left[F(u,v) - \hat{F}(u,v) \right]^2 \right\} &= E \left\{ \left[Q(u,v) + N(u,v) \right]^2 \right\} \\ &= E \left\{ Q^2(u,v) + 2Q(u,v)N(u,v) + N^2(u,v) \right\} \quad (3.30) \\ &= E \left\{ Q^2(u,v) \right\} + 2E \left\{ Q(u,v) \right\} E \left\{ N(u,v) \right\} + E \left\{ N^2(u,v) \right\} \end{aligned}$$

where the independence of Q and N has been used to break the cross term expectation into the product of the expected values of Q and N . Each of the three terms of (3.30) (which will be referred to as the mean square quantization error, mutual error and mean square noise error, respectively) will be evaluated separately in the following three sections.

Mean Square Quantization Error. The $E\{Q^2(u,v)\}$ can be determined

quite easily on the basis of the discussion presented in the previous section on uniform quantization of the DCT coefficients. For a $N(0,1)$ Gaussian random variable the quantization error, which will be called $Q^2[n(u,v)]$, depends solely on the number of bits $n(u,v)$ used to represent the output, as shown in Table III.4. Therefore, the mean square quantization error for the (u,v) DCT coefficient, which is $N[0, \sigma^2(u,v)]$, is given by

$$E\{Q^2(u,v)\} = \sigma^2(u,v) Q^2[n(u,v)] \quad (3.31)$$

Mutual Error. The expected value of the mutual error term is zero. The reason for this is the expected value of the quantization error $\{Q(u,v)\}$ is equal to zero. This can be seen quite easily by analysis of Figure 3.7 which shows the structure of the uniform quantizer.

The $E\{Q(u,v)\}$ is defined by

$$E\{Q(u,v)\} = E\{F(u,v) - F_Q(u,v)\} \quad (3.32)$$

where $F_Q(u,v)$ is the reconstructed value of the quantizer output given that $F(u,v)$ is the input. It has been previously shown that the $E\{F(u,v)\} = 0$. Therefore the $E\{Q(u,v)\}$ equals $-E\{F_Q(u,v)\}$. Figure 3.7 shows that the values $F_Q(u,v)$ can take are symmetrically distributed around 0 as is the probability density of the input to the quantizer. The probability that $F_Q(u,v)$ equals a given value (x) is simply the probability that the input to the quantizer will fall within the bin which has x as an output. Due to the symmetry of the quantizer $P\{F_Q(u,v)=x\} = P\{F_Q(u,v)=-x\}$, hence the expectation of $F_Q(u,v)$ will be 0 as will be $E\{Q(u,v)\}$.

Mean Square Noise Error (Ref 26:23-25). The error due to noise is simply the result of quantization level x_1 being sent and x_k being decoded

in the receiver.

$$E\{N^2(u,v)\} = \sum_{k=0}^{L(u,v)-1} \sum_{l=0}^{L(u,v)-1} (x_k - x_l)^2 P(l,k) \quad (3.33)$$

where

$L(u,v)$ = number of levels assigned to the (u,v) coefficient
 $= 2^{n(u,v)}$ (For compactness $L(u,v)$ will be written
simply as L throughout the remainder of this development)

$P(l,k)$ = Probability that level l was sent and level k received

For the uniform quantizer

$$(x_k - x_l)^2 = d^2(u,v) \sigma^2(u,v) (k-1)^2 \quad (3.34)$$

where

$d(u,v)$ = output level spacing for the (u,v) coefficient
 k, l = number of the output level corresponding to x_k and x_l
respectively

Also, using Bayes Theorem

$$P(l,k) = P(k|l)P(l) \quad (3.35)$$

where

$P(k|l)$ = Probability of level k being received given l was sent
 $P(l)$ = Probability that level l was sent

Substituting equations (3.34) and (3.35) into (3.33) gives

$$E\{N^2(u,v)\} = d^2(u,v) \sigma^2(u,v) \sum_{k=0}^{L-1} \left[\sum_{l=0}^{L-1} (k-1)^2 P(k|l) \right] P(l) \quad (3.36)$$

The term in square brackets can be seen to be the $E\{(k-1)^2/1\}$ which can be expanded as follows:

$$E\{(k-1)^2/1\} = E\{k^2/1\} - 2E\{k/1\} + 1^2 \quad (3.37)$$

The two expected values on the right hand side of equation (3.37) will now be separately evaluated.

As discussed in a previous section the integer k will be expressed in natural binary code for transmission.

$$k = \sum_{i=0}^{n(u,v)-1} k_i 2^i \quad (3.38)$$

where

$$k_i = 0, 1$$

$n(u,v)$ = number of bits assigned to the (u,v) coefficient ($n(u,v)$ will be shortened to n for the remainder of this development)

Therefore

$$E\{k/1\} = \sum_{i=0}^{n-1} E\{k_i/1\} 2^i \quad (3.39)$$

If we assume that the channel is memoryless, the only bit in 1 which impacts the value of k_i is l_i . Further if we assume that the probability of bit i being in error is P_i then

$$\begin{aligned} E\{k_i/1\} &= E\{k_i/l_i\} = (1-P_i)l_i + P_i(1-l_i) \\ &= (1-2P_i)l_i + P_i \end{aligned} \quad (3.40)$$

Substitution of (3.40) into (3.39) gives

$$E\{k/l\} = \sum_{i=0}^{n-1} \{(1-2P_i)l_i + P_i\} 2^i \quad (3.41)$$

$$= \sum_{i=0}^{n-1} l_i 2^i + \sum_{i=0}^{n-1} (1-2l_i) P_i 2^i$$

$$= 1 + \sum_{i=0}^{n-1} (1-2l_i) P_i 2^i$$

Note that the probability of error P_i in equation (3.41) has been allowed to vary for each bit position. This is in anticipation of using different channel coding methods on individual data bits rather than coding the bits all in the same manner. For the specific case where the probability of error for each bit is constant ($P_i = P$ for all i) equation (3.41) reduces to

$$E\{k/l\} = (1-2P) \sum_{i=0}^{n-1} l_i 2^i + P \sum_{i=0}^{n-1} 2^i \quad (3.42)$$

$$= (1-2P)l + P \left[\frac{1-2^n}{1-2} \right]$$

$$= (1-2P)l + P(L-1)$$

where

$$L = L(u, v) = 2^{n(u, v)}$$

Turning now to the calculation of the $E\{k^2/l\}$, using equation (3.38) this expectation can be defined as follows

$$E\{k^2/l\} = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} E\{k_i k_j / l\} 2^{i+j} \quad (3.43)$$

Again assuming that the communication channel is memoryless the values of k_i and k_j depend only on l_i and l_j and in fact $E\{k_i k_j / l\} = E\{k_i / l_i\} E\{k_j / l_j\}$, ($i \neq j$). This point will be demonstrated since it is not an intuitive result.

There are four possible values for the product $k_i k_j$ in terms of l_i and l_j . Assuming that the probability that bit $k_i \neq l_i$ is P_i and the probability that $k_j \neq l_j$ is P_j the four values of $k_i k_j$ and their associated probabilities are

$k_i k_j$	$P(k_i k_j / l_i l_j)$
$l_i l_j$	$(1-P_i)(1-P_j)$
$(1-l_i)l_j$	$(P_i)(1-P_j)$
$l_i(1-l_j)$	$(1-P_i)P_j$
$(1-l_i)(1-l_j)$	$P_i P_j$

Therefore

$$\begin{aligned}
 E\{k_i k_j / l\} &= l_i l_j (1-P_i)(1-P_j) + (1-l_i)l_j P_i (1-P_j) + \\
 &\quad l_i(1-l_j)(1-P_i)P_j + (1-l_i)(1-l_j)P_i P_j \\
 &= [(1-P_i)l_i + P_i(1-l_i)] \cdot [(1-P_j)l_j + P_j(1-l_j)] \\
 &= E\{k_i / l_i\} E\{k_j / l_j\}, \quad (i \neq j)
 \end{aligned} \quad (3.44)$$

For (i=j)

$$E\{k_i k_j / l\} = E\{k_i^2 / l_i\} = (1-P_i)l_i^2 + P_i(1-l_i^2) \quad (3.45)$$

Since l_i can only take on values of 0 or 1, $l_i^2 = l_i$ so

$$\begin{aligned} E\{k_i^2 / l_i\} &= (1-P_i)l_i + P_i(1-l_i) \\ &= (1-2P_i)l_i + P_i \\ &= E\{k_i / l_i\} \end{aligned} \quad (3.46)$$

Substitution of (3.46) and (3.44) into (3.43) gives

$$E\{k^2 / l\} = \sum_{i=0}^{n-1} \left[(1-2P_i)l_i + P_i \right] 2^{2i} + \quad (3.47)$$

$$\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \left[(1-2P_i)l_i + P_i \right] \left[(1-2P_j)l_j + P_j \right] 2^{i+j}$$

$i \neq j$

Adding the missing $i=j$ terms to the double summation and subtracting them from the single sum of equation (3.47) yields

$$E\{k^2 / l\} = \sum_{i=0}^{n-1} \left\{ \left[(1-2P_i)l_i + P_i \right] - \left[(1-2P_i)l_i + P_i \right]^2 \right\} 2^{2i} \quad (3.48)$$

$$+ \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \left[(1-2P_i)P_i + P_i \right] \left[(1-2P_j)l_j + P_j \right] 2^{i+j}$$

The double sum in (3.48) is equal to $E^2\{k/l\}$. Expanding the squared term

in the single sum of (3.48), using the fact that $l_i^2 = l_i$ and cancelling terms gives

$$\begin{aligned} \sum_{i=0}^{n-1} \left\{ \left[(1-2P_i)l_i + P_i \right] - \left[(1-2P_i)l_i + P_i \right]^2 \right\} 2^{2i} \\ = \sum_{i=0}^{n-1} P_i(1-P_i) 2^{2i} \end{aligned} \quad (3.49)$$

Substitution of (3.49) into (3.48) gives

$$\begin{aligned} E\{k^2/l\} &= \sum_{i=0}^{n-1} P_i(1-P_i) 2^{2i} + E^2 \{k/l\} \\ &= \sum_{i=0}^{n-1} P_i(1-P_i) 2^{2i} + \left[1 + \sum_{i=0}^{n-1} (1-2l_i)P_i 2^i \right]^2 \end{aligned} \quad (3.50)$$

For the particular case of $P_i = P$ for all i , equation (3.50) reduces to

$$E\{k^2/l\} = \frac{P(1-P)(L^2-1)}{3} + \left[(1-2P)l + (L-1)P \right]^2 \quad (3.51)$$

Recall that the point of these calculations is to find an expression for $E\{(k-1)^2/l\}$ defined by (3.37) and then in turn to substitute that expression into (3.36) to find a final expression for $E\{N^2(u,v)\}$. Substitution of (3.50) and (3.41) into (3.37) gives

$$\begin{aligned} E\{(k-1)^2/l\} &= \sum_{i=0}^{n-1} P_i(1-P_i) 2^{2i} + \left[1 + \sum_{i=0}^{n-1} (1-2l_i)P_i 2^i \right]^2 \\ &\quad - 2l \left[1 + \sum_{i=0}^{n-1} (1-2P_i)P_i 2^i \right] + l^2 \end{aligned} \quad (3.52)$$

After some algebra this reduces to

$$E\{(k-1)^2/l\} = \sum_{i=0}^{n-1} P_i (1-P_i) 2^{2i} + \left[\sum_{i=0}^{n-1} (1-2l_i) P_i 2^{2i} \right]^2 \quad (3.53)$$

Substitution of (3.53) into (3.36) and expansion of the squared term in (3.53) into a double summation gives

$$E\{N^2(u,v)\} = d^2(u,v) \sigma^2(u,v) \left\{ \sum_{l=0}^{L-1} \left[\sum_{i=0}^{n-1} P_i (1-P_i) 2^{2i} + \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (1-2l_i) P_i (1-2l_j) P_j 2^{i+j} \right] P(l) \right\} \quad (3.54)$$

Since the first sum in the square brackets does not depend on l it can be pulled outside the summation. Then noting that $\sum_{l=0}^{L-1} P(l) = 1$, equation (3.54) becomes

$$E\{N^2(u,v)\} = d^2(u,v) \sigma^2(u,v) \left\{ \sum_{i=0}^{n-1} P_i (1-P_i) 2^{2i} + \sum_{l=0}^{L-1} \left[\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (1-2l_i) P_i (1-2l_j) P_j 2^{i+j} \right] P(l) \right\} \quad (3.55)$$

At this point it is necessary to define $P(l)$. This is simply the probability that the input (x) to the quantizer, which is a $N(0,1)$ random variable, falls into the l th quantizer level. Reviewing Figure 3.7 it can be seen that

$$P(l) = P \left\{ \left[l - \frac{L}{2} \right] d(u,v) \leq x \leq \left[(l+1) - \frac{L}{2} \right] d(u,v) \right\}, \quad l = 1, 2, \dots, L-2 \quad (3.56)$$

$$P(l) = \begin{cases} P \left[-\infty \leq x \leq (1 - \frac{L}{2})d(u, v) \right], & l = 0 \\ P \left[(\frac{L}{2} - 1)d(u, v) \leq x \leq \infty \right], & l = L-1 \end{cases} \quad (3.56)$$

Since x is $N(0,1)$, $P(l)$ can be defined in terms of the error function, $\text{erf}(x)$.

Defining $\text{erf}(x)$ as follows

$$\text{erf}(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \quad (3.57)$$

the equations for $P(l)$ become

$$P(l) = \begin{cases} \text{erf} \left[(1+l - \frac{L}{2})d(u, v) \right] - \text{erf} \left[(1 - \frac{L}{2})d(u, v) \right], & l = 1, 2, \dots L-2 \\ \text{erf} \left[(1 - \frac{L}{2})d(u, v) \right] & , l = 0 \\ 1 - \text{erf} \left[(\frac{L}{2} - 1)d(u, v) \right] & , l = L-1 \end{cases} \quad (3.58)$$

Due to the complexity of the expression for $P(l)$, there is no way to simplify equation (3.55) any further so it will be taken as the final expression for $E\{N^2(u, v)\}$ where the probability of a bit error (P_i) depends on the bit position. If P_i is constant for all i (3.55) can be simplified considerably. Setting $P_i = P$ for all i in equation (3.55) gives

$$E\{N^2(u, v)\} = d^2(u, v) \sigma^2(u, v) \left\{ P(1-P) \sum_{i=0}^{n-1} 2^{2i} + \right. \quad (3.59)$$

$$\left. P^2 \sum_{l=0}^{L-1} \left[\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (1-2l_i)(1-2l_j) 2^{i+j} \right] P(l) \right\}$$

Noticing that

$$\sum_{i=0}^{n-1} (2^2)^i = \frac{1 - 2^{2n}}{1 - 2^2} = \frac{2^{2n} - 1}{3} = \frac{L^2 - 1}{3} \quad (3.60a)$$

and

$$\begin{aligned} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (1-2l_i)(1-2l_j)2^{i+j} &= \sum_{i=0}^{n-1} (1-2l_i)2^i \sum_{j=0}^{n-1} (1-2l_j)2^j \\ &= \left[\sum_{i=0}^{n-1} 2^{i-2} \sum_{i=0}^{n-1} l_i 2^i \right] \left[\sum_{j=0}^{n-1} 2^{j-2} \sum_{j=0}^{n-1} l_j 2^j \right] \quad (3.60b) \\ &= \left[2^{n-1-2l} \right]^2 \\ &= \left[L-1-2l \right]^2 \end{aligned}$$

equation (3.59) becomes

$$E\{N^2(u,v)\} = d^2(u,v) \sigma^2(u,v) \left\{ \frac{P(1-P)(L^2-1)}{3} + P^2 \sum_{l=0}^{L-1} \left[L-1-2l \right]^2 P(l) \right\} \quad (3.61)$$

Expansion of the squared term in the summation of (3.61) gives

$$\begin{aligned} \sum_{l=0}^{L-1} (L-1-2l)^2 P(l) &= \sum_{l=0}^{L-1} \left[(L-1)^2 - 4l(L-1) + 4l^2 \right] P(l) \quad (3.62) \\ &= (L-1)^2 \sum_{l=0}^{L-1} P(l) - 4(L-1) \sum_{l=0}^{L-1} l P(l) + 4 \sum_{l=0}^{L-1} l^2 P(l) \end{aligned}$$

The term $\sum_{l=0}^{L-1} P(l)$ equals 1. The $\sum_{l=0}^{L-1} l P(l)$ is just equal to $E[l]$. Since the input to the quantizer is symmetric about 0 and the quantizer is also

symmetric about 0 the expected value of the output of the quantizer is the midpoint of the range of output values.

$$E\{l\} = \frac{L-1}{2} \quad (3.63)$$

Finally the $\sum_{l=0}^{L-1} l^2 P(l)$ is simply $E\{l^2\}$. No simple expression for the expected value of l^2 can be given since $P(l)$ is not constant for all values of l . However $E\{l^2\}$ can be tabulated as a function of $n(u,v)$, the number of bits assigned to a coefficient, as shown in Table III.5. $E\{l^2\}$ will simply be called $SQ(u,v)$ in the remaining equations.

Substituting the expressions for $E\{l\}$ and $E\{l^2\}$ into (3.62) yields

$$\begin{aligned} \sum_{l=0}^{L-1} (L-1-2l)^2 P(l) &= (L-1)^2 - \frac{4(L-1)(L-1)}{2} + 4SQ(u,v) \\ &= 4SQ(u,v) - (L-1)^2 \end{aligned} \quad (3.64)$$

Then substituting (3.64) into (3.61) gives the final expression for $E\{N^2(u,v)\}$ for the constant probability of error channel.

$$E\{N^2(u,v)\} = d^2(u,v) \sigma^2(u,v) \left\{ \frac{P(1-P)(L^2-1)}{3} + P^2 \left[4SQ(u,v) - (L-1)^2 \right] \right\} \quad (3.65)$$

Total Error. Taking the expressions for $E\{Q^2(u,v)\}$ and $E\{N^2(u,v)\}$ found in the previous sections and summing them gives the total distortion (D_T). If the probability of an error occurring in the i th bit position is P_i , then

$$D_T = \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} \frac{A^2(u,v)}{4} \sigma^2(u,v) \left\{ Q^2 \left[n(u,v) \right] + d^2(u,v) \left(\sum_{i=0}^{n(u,v)-1} P_i (1-P_i) 2^{2i} + \right. \right. \quad (3.66)$$

$$\left. \left. \sum_{i=0}^{L(u,v)-1} \left[\sum_{j=0}^{n(u,v)-1} \sum_{j=0}^{n(u,v)-1} (1-2l_i) P_i (1-2l_j) P_j 2^{i+j} \right] P(l) \right\} \right\}$$

Table III.5

Expected Value of the Square of
the Uniform Quantizer Output Level

<u>Number of Bits</u>	<u>Number of Levels</u>	<u>$E\{l^2\}$</u>
1	2	.50000000
2	4	3.1388048
3	8	15.052876
4	16	65.047276
5	32	268.40287
6	64	1084.4977
7	128	4341.3770
8	256	17312.877
9	512	68953.721
10	1024	274588.16
11	2048	1093803.7
12	4096	4359072.6
13	8192	17380072.
14	16384	69325881.
15	32768	.27663295E+09
16	65536	.11041433E+10

For $P_i = P$ for all i ,

$$D_T = \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} \frac{A_p^2(u,v)}{4} \sigma^2(u,v) \left\{ Q^2[n(u,v)] + \right. \quad (3.67)$$

$$\left. d^2(u,v) \left(\frac{P(1-P)}{3} (L^2(u,v)-1) + P^2 [4SQ(u,v) - (L(u,v)-1)^2] \right) \right\}$$

where $Q^2[n(u,v)]$, $SQ(u,v)$ and $d(u,v)$ are tabulated functions which depend on $n(u,v)$ and $L(u,v) = 2^{n(u,v)}$. Thus given $n(u,v)$ and P , D_T can be computed using equation (3.67). Use of equation (3.66) will be deferred until the next chapter when error correction coding is applied.

Calculated Performance of the Source Encoded Image Transmission System

Using equations (3.24) and (3.67) (constant P) the performance of the source encoded imagery transmission system was calculated by computer. The results are displayed in Figure 3.8. The graph plots the average number of bits per sample (R) versus the calculated distortion (D) for that number of bits. That is

$$R = \frac{1}{M^2} \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} n(u,v) \quad (3.68)$$

$$D = D_T [n(u,v), P]$$

The calculation was performed for five different values of $P(0, 10^{-6}, 10^{-5}, 10^{-4}$ and 10^{-3}). The power-normalized version of the weighting function $A_p(u,v)$ was used in the calculation. Also displayed in Figure 3.8 is the approximate rate distortion function for the DCT coefficients.

The first conclusion which can be drawn from Figure 3.8 is that the cosine transform/block quantization (DCT/Q) source encoding scheme

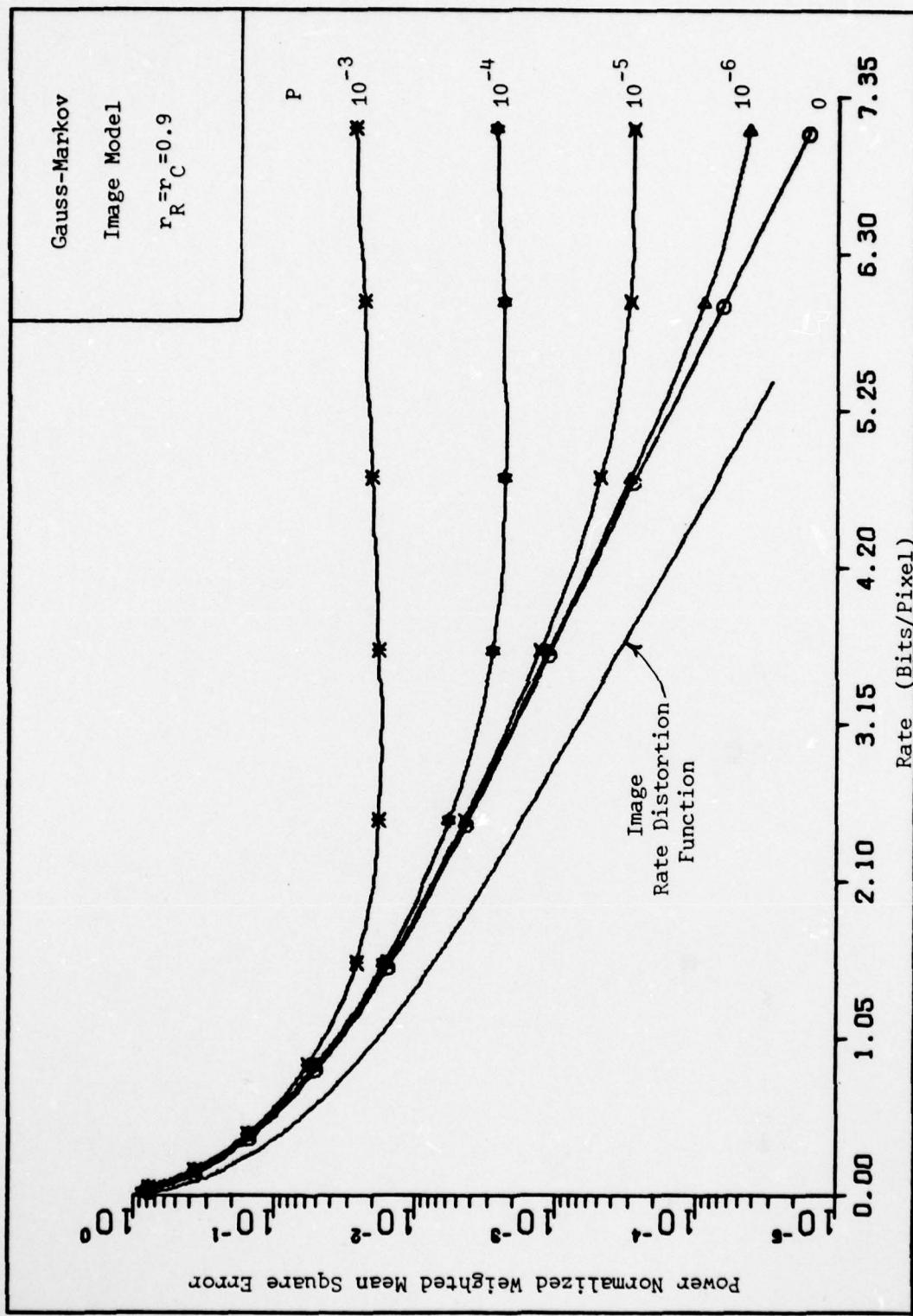


Figure 3.8 Performance of the Source Encoded Imagery Transmission System

is quite efficient as long as the communications channel noise is low. For $P \leq 10^{-6}$ the performance of the source encoder is within 1 bit/pixel of the rate distortion bound.

The performance of the source encoder deteriorates, however, as the probability of error increases. Figure 3.8 shows that for bit error rates in the range of 10^{-3} (which is typical for many standard military communications systems), the DCT/Q technique can not reach distortion levels less than 1% no matter how many bits are transmitted. In fact the curves at these bit error rates exhibit the characteristic that beyond a certain rate (approximately 2 bits/pixel for $P = 10^{-3}$) the distortion actually gets worse if more source bits are added. This counter intuitive result can be explained through an analysis of the components of the distortion.

Table III.6 shows the calculated values of D and R that are plotted in Figure 3.8 for $P = 10^{-3}$. The quantization and noise components of the distortion are also separately broken out in Table III.6. The table shows that the quantization error is inversely related to the rate as expected. On the other hand, the noise error increases as the rate increases. This result is also to be expected. As the rate increases, more bits are transmitted per pixel so naturally more errors will occur. More importantly, as R goes up the number of quantization levels assigned to each DCT coefficient increases. The binary coding structure used to represent the quantizer output causes some bits in the transmitted data stream to be much more important than others. If an error occurs in one of these bits, a significant amount of distortion results. Increasing the number of quantization levels only worsens this situation. Therefore the noise

Table III.6
Performance of the Source Encoder for $P=10^{-3}$

R	D	Quantization Error	Noise Error
6.03515625	.02402693	.00005913	.02396780
5.93750000	.02354282	.00006710	.02347572
5.79687500	.02340024	.00007942	.02332082
5.66796875	.02277700	.00009359	.02268341
5.58984375	.02275597	.00010241	.02265355
5.44921875	.02256076	.00012051	.02244024
5.38421875	.02205320	.00014092	.02191228
5.21093750	.02133659	.00016271	.02117389
5.10546875	.02118413	.00018436	.02099978
4.97656250	.02074980	.00021748	.02053232
4.87109375	.02066659	.00024531	.02042128
4.75390625	.02055927	.00028227	.02027699
4.62890625	.02002959	.00032888	.01970071
4.50390625	.01987773	.00038138	.01949634
4.42187500	.01983159	.00042123	.01941036
4.29296875	.01945165	.00048600	.01896564
4.17968750	.01878605	.00056212	.01822393
4.05078125	.01872053	.00065431	.01806622
3.93750000	.01834864	.00075398	.01759465
3.81250000	.01833547	.00087028	.01746520
3.71484375	.01834768	.00097066	.01737703
3.60546875	.01794601	.00111302	.01683300
3.47265625	.01793267	.00129864	.01663403
3.34375000	.01767256	.00151042	.01616214
3.24609375	.01717923	.00172320	.01545604
3.13671875	.01723763	.00194507	.01529256
3.00000000	.01743287	.00225800	.01517488
2.91796875	.01731959	.00250191	.01481768
2.78515625	.01759919	.00291306	.01468613
2.67187500	.01748373	.00336647	.01411726
2.57812500	.01778436	.00376152	.01402283
2.46484375	.01816715	.00425956	.01390759
2.34765625	.01837497	.00493617	.01343980
2.25390625	.01843037	.00567442	.01275595
2.14843750	.01902214	.00639584	.01262630
2.03515625	.01968998	.00749706	.01219292
1.94531250	.02048301	.00835917	.01212384
1.85156250	.02145458	.00945281	.01200177
1.75781250	.02230181	.01078915	.01151266
1.66406250	.02360017	.01225788	.01134229
1.59375000	.02460231	.01367077	.01093155
1.50781250	.02592267	.01553591	.01038676
1.42187500	.02798305	.01783971	.01014334
1.34375000	.02990278	.01979505	.01010772
1.28515625	.03160685	.02187811	.00972874
1.19140625	.03486533	.02525432	.00961101
1.12109375	.03811768	.02900577	.00911190
1.06640625	.04020455	.03111627	.00908828
1.00390625	.04383073	.03490270	.00892804

error is directly correlated to the rate. For $P = 10^{-3}$, Table III.6 shows that the noise term becomes dominant for $R > 2$ so the total distortion actually increases above this rate.

The final conclusion which Figure 3.8 yields is that there is definitely room for a trade off of source bits for channel encoding bits in the DCT/Q system. Since adding source bits beyond a certain level will not improve the DCT/Q performance and may actually worsen it, a reasonable alternative would be to add error correction coding bits instead to reduce the probability of bit error. This trade off of source data bits and channel encoding bits is the subject of the next chapter.

IV. Performance Improvement With Combined Source/Channel Encoding Procedures

In this chapter, the effect of adding error correction coding bits to the output of the source encoder will be calculated. No attempt will be made to produce an optimum source/channel encoding procedure. The emphasis will be on applying standard error correction procedures in order to reduce the overall distortion level in the received image.

Channel Coding Procedures

There are two broad categories of error correcting codes, block codes and convolutional codes. Block codes accept a fixed-length block of K data bits, and output an N bit codeword ($N > K$) which depends only on those K information bits. Convolutional codes also operate on K bits of data at a time, but the N bit codeword which they produce depends not only on the K information bits, but also on M previous data bits. Both types of codes are normally identified by an (N, K) descriptor. Because some indication of the size of the memory, M , is needed for convolutional codes, they normally have a third identifying number, known as the constraint length (C), associated with them. The definition of constraint length varies somewhat in the literature, but for this report, C will be defined to be the sum of K and M .

Since both types of codes transmit N bits for all K bits of data, the effective transmission rate R_e over the communications channel is N/K times the data rate of the source. The ratio N/K is called the rate of the code and has units of channel bits/source bit.

The final characteristic of channel codes, which needs to be defined, is a performance measure. The standard measure used in much of the research

literature is P_E , the probability of a codeword being decoded incorrectly.

For the purpose of calculating the performance of the source/channel encoder, the error rate that is needed is P_e , the probability of an individual data bit being received in error. There is, in general, no direct relationship between P_E and P_e . In fact, the calculation of P_E for most codes is extremely difficult, if not impossible. Much research has focused on simply finding tight upper and lower bounds for P_E . For this reason one of the primary criteria used to select the channel codes was the availability of calculated or empirically derived values for P_e .

Block Codes. The structure of block codes has been studied extensively, and as a result, there are literally thousands of potential codes from which to choose. The block codes chosen for application to the imagery data link were selected because they are "perfect" codes. A perfect code is one that will correct all n and fewer bit errors in a codeword ($n=1, 2, 3 \dots$) but will not correct any other error patterns. In this sense perfect codes are the most efficient codes for correcting such error patterns. That is why perfect codes are among the most thoroughly studied and often-implemented block codes. There are three families of perfect codes. One code from each family was selected for study.

The first family consists of the $(N,1)$ repetition codes where N is an odd number. For this type of code, there are only two codewords, all 1's or all 0's. The codeword is decoded by majority vote. The $(N,1)$ repetition code will correct all $(N-1)/2$ and fewer bit errors in the received codeword. Because only 1 bit of information is transmitted in each codeword, the shortest repetition code, $(3,1)$, was the one selected from this group.

The $(3,1)$ repetition code corrects all single bit errors. Therefore,

the probability of the data bit being received in error, P_e , is simply the probability that two or three errors occur in the codeword. For the memoryless, binary symmetric communications channel (BSC) with bit error probability, P , the performance of the (3,1) code is given by

$$\begin{aligned} P_e &= 3P^2(1 - P) + P^3 \\ &= 3P^2 - 2P^3 \end{aligned} \quad (4.1)$$

The (3,1) repetition code is also a member of the second class of perfect codes, the Hamming codes. These codes correct all single bit errors and have the following specific relationship between N and K

$$\left. \begin{aligned} K &= 2^i - 1 - i \\ N &= 2^i - 1 \end{aligned} \right\} \quad i = 0, 1, 2, 3 \dots \quad (4.2)$$

The (3,1) code has the worst rate of all the Hamming codes. The (7,4) Hamming code was selected for use in the imagery transmission system since its performance is more representative of the overall Hamming family performance than the (3,1) code.

Since the (7,4) Hamming code corrects all single bit errors in the codewords, but no more, the probability of a codeword being decoded incorrectly is simply the probability of two or more bits being in error. For the memoryless BSC

$$P_E = \sum_{i=2}^7 \binom{7}{i} P^i (1 - P)^{7-i} \quad (4.3)$$

However, even if a codeword is in error, some of the individual bits will still be decoded correctly. McEliece shows that the probability of a bit

error, P_e , for the (7,4) Hamming code is given by (Ref 25:7)

$$P_e = 9P^2(1-P^5) + 19P^3(1-P)^4 + 16P^4(1-P^3) + \quad (4.4)$$

$$12P^5(1-P)^2 + 7P^6(1-P) + P^7$$

$$= 9P^2 - 26P^3 + \text{higher order terms}$$

For $P \leq 0.001$, the higher order terms in (4.4) are insignificant.

The final block code selected is the (23,12) Golay code. The (23,12) Golay code corrects all 3 or fewer bit errors and is a very powerful channel encoding technique. Rather than calculate P_E directly for the Golay code, it is easier to calculate the probability of a codeword being received correctly, P_C , and then find P_E by $P_E = 1 - P_C$. For the (23,12) code

$$P_C = (1-P)^{23} + \binom{23}{1} (1-P)^{22}P + \binom{23}{2} (1-P)^{21}P^2 + \quad (4.5)$$

$$\binom{23}{3} (1-P)^{20}P^3$$

For $P = 10^{-3}$, P_C is equal to .99999999913 or $P_E = 8.7 \times 10^{-9}$. This means that out of every 115 million codewords sent, one will be decoded incorrectly. The probability of any one data bit being in error is still lower, because even in the one block received in error, some of the bits will be decoded correctly. No specific formula was found for the probability of bit error for the Golay code so the probability of block error will be used in the performance calculation. The error rate is so close to zero that this approximation will have little or no impact on the results.

Convolutional Codes. The technology associated with convolutional codes is not nearly as well developed as block code technology. Only in

recent years have efficient decoding procedures been developed for them. One of the factors which greatly complicates the systematic study of convolutional codes is the constraint length, C . As an example, a (2,1) block code is of no practical use since it has no error correcting capability. A (2,1) convolutional code, on the other hand, is practical. In fact, the study of (2,1) codes with variations in C is a whole subject in and of itself.

Despite the difficulties involved with the development of convolutional codes, they have been found to provide significant performance improvement over block codes in high noise ($P > 10^{-3}$) applications such as communications with space vehicles. The three convolutional codes selected for application to the image transmission system are all products of space programs.

The first of the three convolutional codes has been used on Mariner class spacecraft since 1977. It is a (2,1) code with constraint length 6. This is a very short constraint length code which permits some of the most powerful decoding techniques to be applied, such as the Viterbi decoding rules (Ref 42:237-334). The measured performance of the (2,1) Mariner code is shown in Figure 4.1.

The other two convolutional codes are products of communications satellite technology. One is a (4,3) $C = 80$ code which is used on the INTELSAT communications satellites. It is designed to transmit 40.8 Kbit/sec data over regular INTELSAT voice channels ($P=10^{-4}$) at an effective bit error rate of 10^{-9} (Ref 45:942). This code will be referred to as the (4,3) SPADE code since it is incorporated in the SPADE system of the INTELSAT network. The full measured performance curve of the (4,3) SPADE code is shown in Figure 4.2.

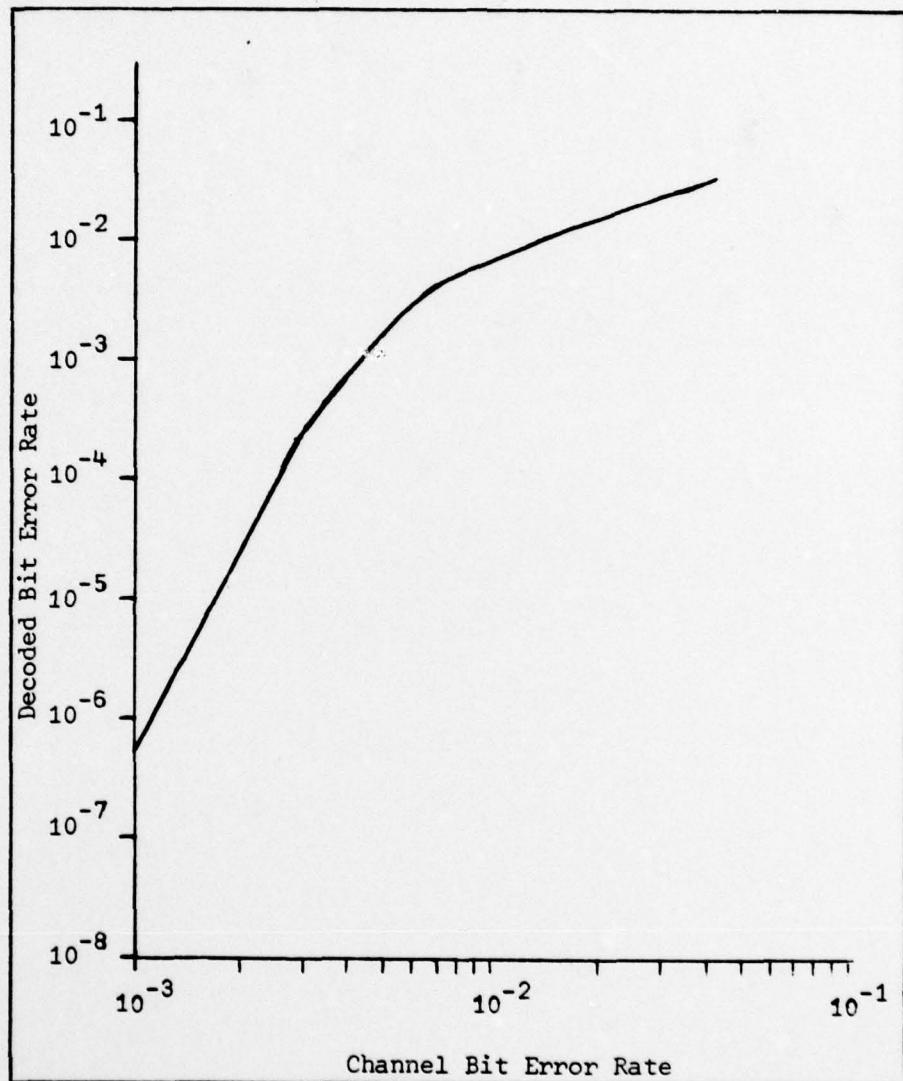


Figure 4.1 Performance of (2,1) Mariner Code

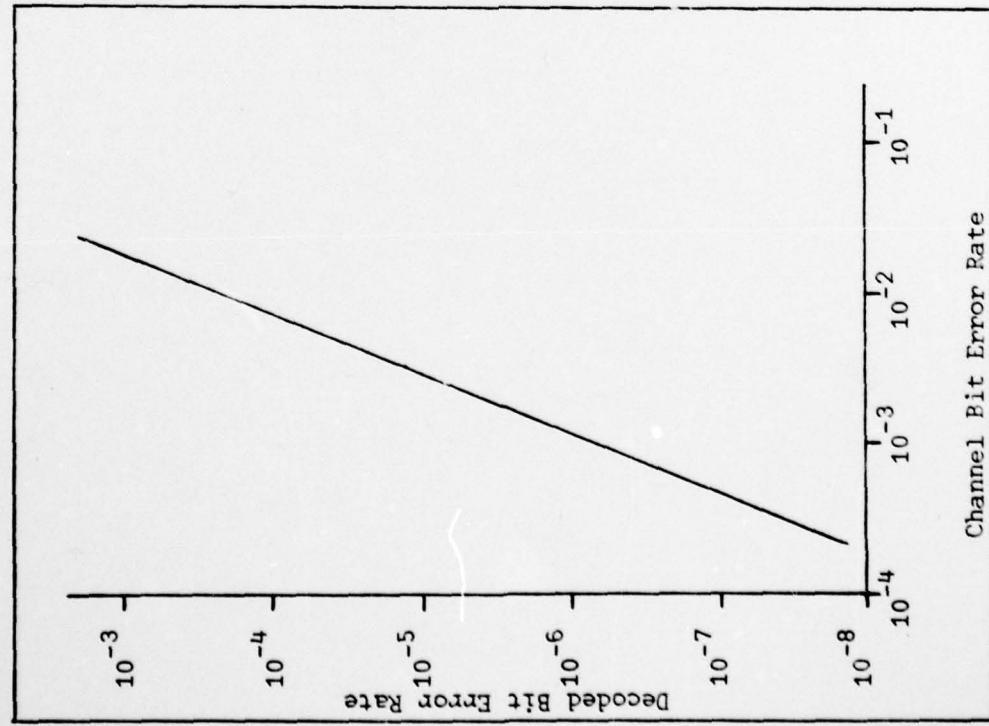


Figure 4.2 Performance of (4,3) SPADE Code

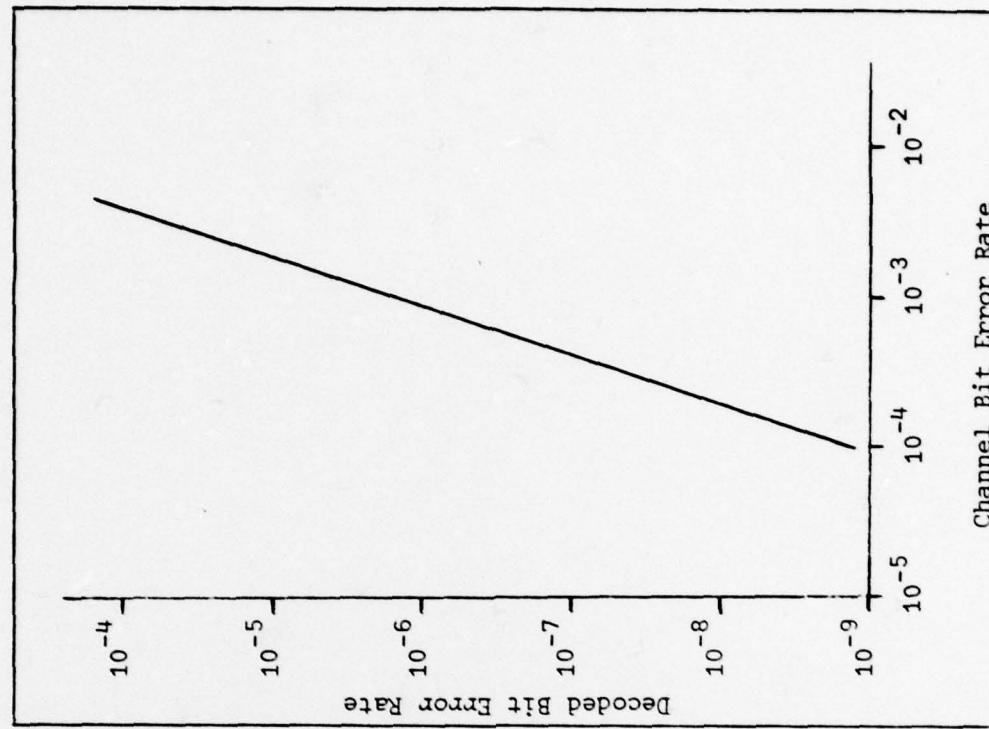


Figure 4.3 Performance of (8,7) DITEC Code

The final convolutional code investigated for this thesis is the (8,7) $C = 1176$ DITEC code. DITEC is a digital television system designed to transmit 525 line color TV signals at a rate of 33.6 Mbit/sec. The (8,7) convolutional code provides a decoder error rate of 10^{-9} on a channel with $P = 10^{-4}$ (Ref 45:943). The full performance curve of the DITEC coder is shown in Figure 4.3.

Summary of Selected Error Correcting Codes. In order to restrict the number of combinations of variables in the source/channel performance calculations, all results were derived assuming a memoryless binary symmetric channel with probability of bit error equal to 10^{-3} . Table IV.1 summarizes the parameters and performance of the six channel codes at this error rate.

Application of Channel Encoding to the Block Quantizer Output

The next step in the process of defining the entire image transmission system is to identify how the channel codes will be applied to the quantizer's output. That is, which bits will be protected against channel errors?

The first obvious answer to this question is to apply the channel codes uniformly to all source bits. This is certainly the most straightforward scheme and would be the easiest to implement. However, the justification for applying the codes in this manner is that all data bits are equally important and should receive equal protection. This is not the case for the quantizer's output. While it is true that all of the variables being quantized have been normalized to be $N(0,1)$, at the receiver the reconstructed values will be rescaled to their original $N\{0,\sigma(u,v)\}$ distributions. By reviewing the table of DCT variances (Table III.2), it can be

AD-A080 172 AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH SCHO0--ETC F/G 9/4
PERFORMANCE OF A SOURCE/CHANNEL ENCODED IMAGERY TRANSMISSION SY--ETC(U)
DEC 79 R A DURYEA
UNCLASSIFIED AFIT/GE/EE/79D-12

NL

2 OF 2
AD
A080172



END
DATE
FILED
3-80
DDC

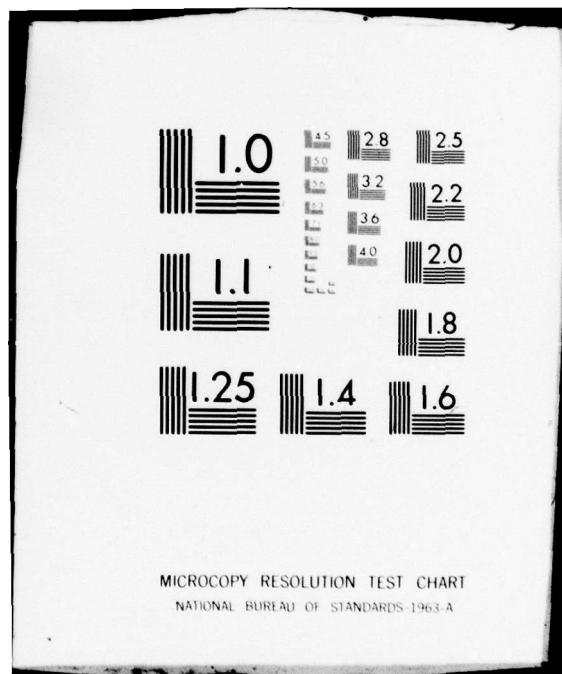


Table IV.1
Summary of Channel Encoder Performance at $P = 10^{-3}$

<u>Code</u>	Codeword Length <u>N</u>	Data Bits/ Codeword <u>K</u>	Constraint Length <u>C</u>	Error Rate P_e	Code Rate <u>N/K</u>
Repetition	3	1	Block	3.0×10^{-6}	3.00
Hamming	7	4	Block	9.0×10^{-6}	1.75
Golay	23	12	Block	$<8.7 \times 10^{-9}$	1.92
Mariner	2	1	6	5.5×10^{-7}	2.00
SPADE	4	3	80	8.5×10^{-7}	1.33
DITEC	8	7	1176	3.5×10^{-6}	1.14

seen that the effect of the rescaling will be to greatly deemphasize the effect of transmission errors in the high frequency components. Errors in the low frequency components will be passed through relatively unchanged or perhaps even magnified.

In order to take into account the variability of the impact of transmission errors, a reasonable second cut at application of channel coding would be to protect only a specific subset of DCT coefficients. The questions then becomes, what subset should be protected? For simplicity of implementation, the easiest subset to select is an $m \times m$ block of the low frequency coefficients. However, a review of the bit assignment matrices in Table III.3, indicates that a better choice might be a roughly triangular shaped block of low frequency components or all the coefficients in the first several rows and columns.

Protecting subsets of the DCT coefficients still involves the assumption that the individual bits within the subset all carry the same amount

of information. Again, this is not the case for the quantizer's output. Since the output of the quantizer consists of natural binary code words, the most significant bits of those words carry more information than the other bits. Therefore, a third alternative in applying error correcting codes would be to protect the most significant bits of each quantizer word and leave the other bits unprotected. This application was anticipated in the noise error calculation of Chapter III, by allowing the probability of a data bit error to be dependent on the bit's position (Equation 3.66). It must be pointed out that a scheme to protect individual bits would be difficult to implement. The only reasonable codes which could be applied are the (N,1) repetition block codes or very short constraint length (N,1) convolutional codes.

Theoretical performance curves were calculated for all three channel coding procedures proposed above: uniform coding of all bits, coding of selected subsets of DCT coefficients and coding of the most significant bits of each quantizer word.

Calculated Performance of the Source/Channel Encoded Imagery Transmission System

Uniform Application of Error Correcting Codes. In order to calculate the performance of the imagery transmission system with all data bits channel encoded uniformly with an (N,K) code, equations (3.67) and (3.68) were modified as follows:

$$R = \frac{N}{KM^2} \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} n(u,v) \quad (4.6)$$

$$D_T = \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} \frac{A^2(u,v)}{4} \sigma^2(u,v) \left\{ Q^2[n(u,v)] + \right. \quad (4.6)$$

$$\left. d^2(u,v) \left(\frac{P_e(1-P_e)\{L^2(u,v)-1\}}{3} + P_e^2 \left\{ 4SQ(u,v) - [L(u,v)-1]^2 \right\} \right) \right\}$$

where P_e is the probability of a decoded bit being in error.

The performance curves for the block codes are shown in Figure 4.4 and for the convolutional codes in Figure 4.5. Also shown in the two figures is the calculated performance of the source encoder alone at a bit error probability of 10^{-3} . The graphs clearly show that the performance of the imagery transmission system can be improved significantly through the application of error correction codes.

For the three block codes, the lowest distortion levels were achieved by the (7,4) Hamming code. This is somewhat counterintuitive, since the (23,12) Golay code has a much lower P_e . However, both codes provide a $P_e \leq 10^{-5}$. For error rates in this range, the overall distortion is dominated by the quantization error, which is not dependent in any way on the channel code. As a result, the difference in performance of the two codes lies in their code rates. The (7,4) code has a slightly lower rate (1.75 channel bits/source bit) than the (23,12) code (1.92 channel bits/source bit), so the Hamming code requires fewer bits to achieve the same level of distortion as the Golay code. For similar reasons, the (3,1) repetition code's performance is inferior to that of the other codes.

The convolutional codes showed similar results. Both the (4,3) SPADE code and (8,7) DITEC code have lower N/K ratios than the block codes, and both provided better performance. For the (7,4) Hamming code, the rate at which the source/channel coder's performance matches the source encoder's

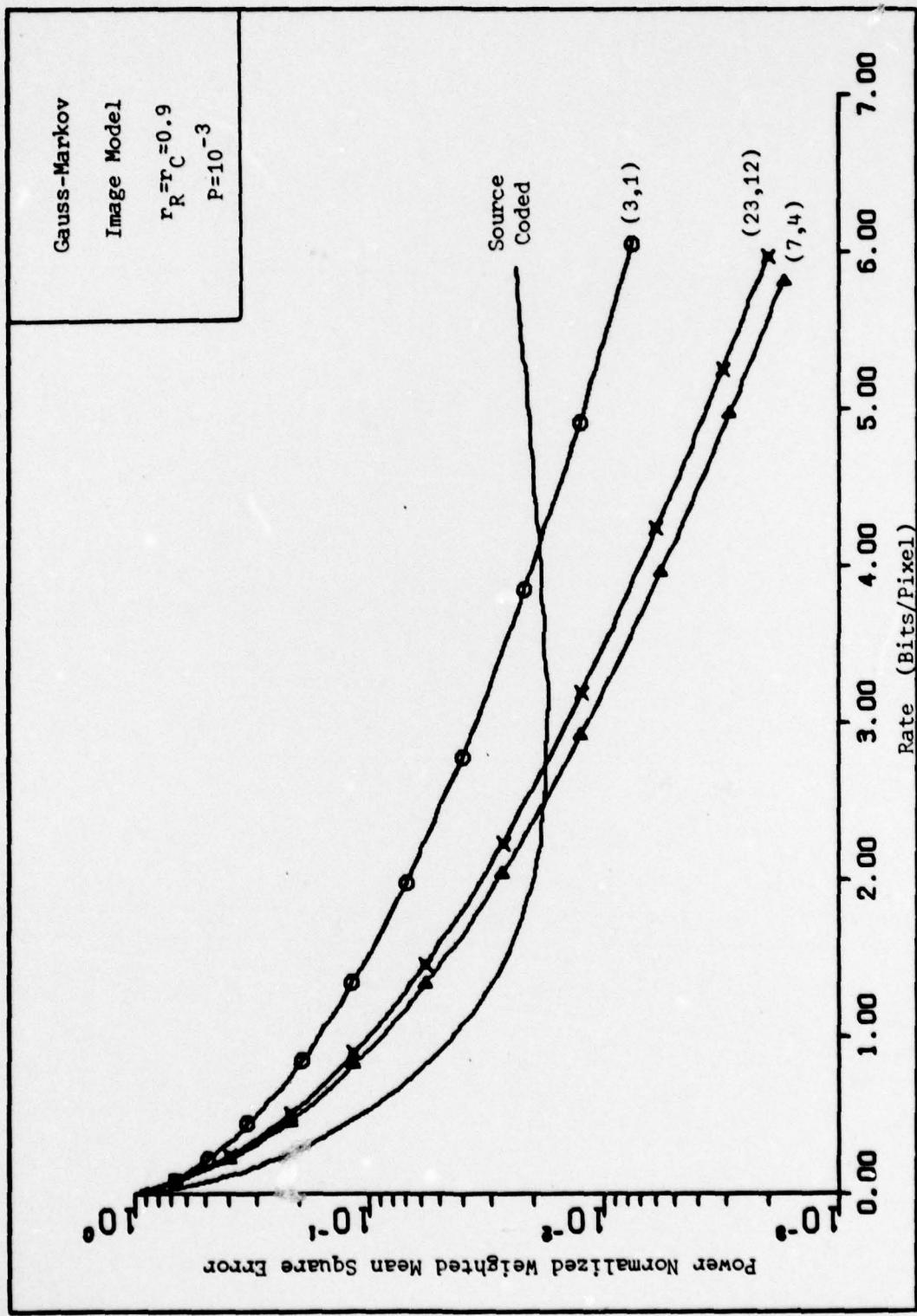


Figure 4.4 Performance of Source/Channel Encoded Imagery Transmission System
 Using (N, K) Block Codes

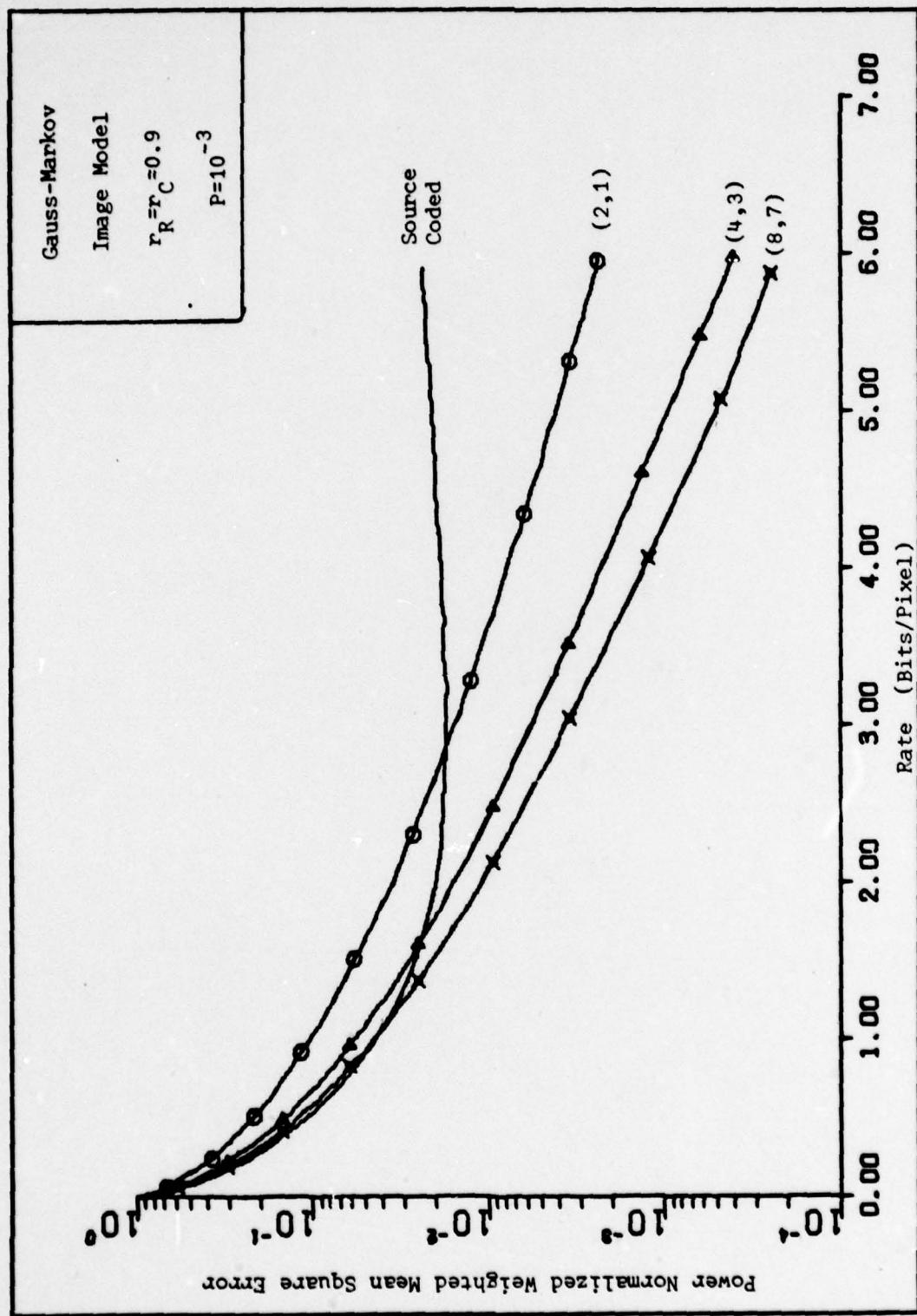


Figure 4.5 Performance of Source/Channel Encoded Imagery Transmission System
 Using (N, K) Convolutional Codes

performance (this will be called the break even rate) was approximately 2.5 bits/pixel. For the (4,3) and (8,7) convolutional codes, the break even rate has been reduced to 1.6 and 1.1 bits/pixel, respectively.

One additional point must be emphasized about Figures 4.4 and 4.5 and all following graphs in this chapter. In order to present the data in a reasonable format, the distortion levels are plotted on a logarithmic scale. Therefore, the curves tend to visually underestimate the performance improvement achieved by the application of channel codes. As an example of this, the distortion level provided by the (7,4) Hamming code at 3 bits/pixel is only 1/3 that of the uncoded source. For the (8,7) DITEC code, the weighted mean square error at this rate is five times smaller than the source encoder's performance. The application of error correcting codes can provide a very significant performance gain.

Application of Error Correcting Codes to Specific Subsets of DCT Coefficients. The next set of results are intended to show the effect of channel coding only certain subsets of the DCT coefficients. Again, it was necessary to limit the number of varying parameters in the calculations. Therefore, the analysis was performed only for the best block code $[(7,4) \text{ Hamming}]$ and the best convolutional code $[(8,7) \text{ DITEC}]$ as determined by the uniform application results.

The equations used for these calculations are:

$$R = \frac{1}{M^2} \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} c_r(u,v) n(u,v) \quad (4.7)$$

where

$$c_r(u,v) = \begin{cases} N/K \text{ for } (u,v) \text{ in the selected subset} \\ 1 \text{ for } (u,v) \text{ not in the subset} \end{cases}$$

$$D_T = \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} \frac{A^2(u,v)}{4} \sigma^2(u,v) \left\{ Q^2 [n(u,v)] + d^2(u,v) \left(\frac{P(u,v)\{1-P(u,v)\}}{3} \frac{\{L^2(u,v)-1\}}{P^2(u,v) \{4SQ(u,v)-[L(u,v)-1]^2\}} \right) \right\} \quad (4.8)$$

where

$$P(u,v) = \begin{cases} P_e & \text{for } (u,v) \text{ in the selected subset} \\ 10^{-3} & \text{for } (u,v) \text{ not in the subset} \end{cases}$$

The first subsets selected were $m \times m$ square blocks of the lowest frequency components. The results for $m=4, 6$ and 8 are shown in Figures 4.6 and 4.7. For comparison purposes, the performance results for the source coder and the uniformly applied channel code are also included in these figures.

The curves for the $(7,4)$ Hamming code (Figure 4.6) show that protecting only the square subset of samples results in reduced distortion levels for rates between 1.2 and 5.2 bits/pixel. Below this range, the source coder alone achieves the best performance. Above this range, channel coding the entire set of DCT coefficients provides the lowest weighted MSE. Within the range, the subset which provides the lowest distortion depends on the rate. Up to 2 bits/pixel, the $m=4$ subset has the lowest weighted MSE. From 2 bits/pixel to 3 bits/pixel, the $m=6$ curve is the lowest. The $m=8$ performance surpasses the $m=6$ performance at 3 bits/pixel. This pattern of performance curve crossovers is repeated in all $(7,4)$ code results and repeated to a much smaller degree in the $(8,7)$ DITEC code results. For $R=3$, coding only the 6×6 block of

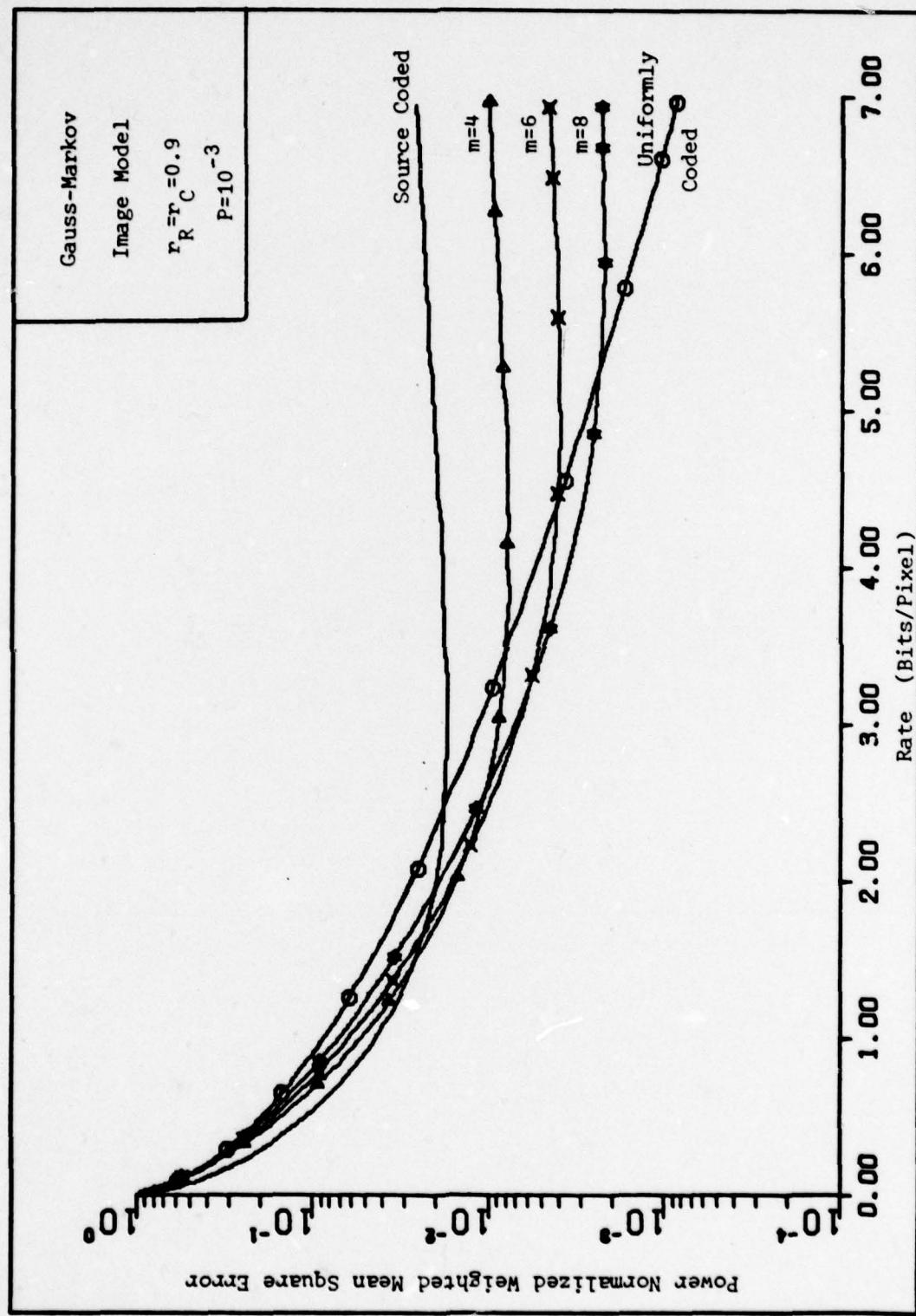


Figure 4.6 Performance of Source/Channel Encoded Imagery Transmission System with an $m \times m$ Square Block of Low Frequency DCT Coefficients $(7,4)$ Hamming Coded

lowest frequency DCT coefficients yields a calculated WMSE of 0.0068.

This is 40% less than the comparable figure for the uniformly coded case (0.011) and 60% less than the source coder's performance (0.017).

Figure 4.7 shows the calculated results of coding an $m \times m$ square block with the (8,7) DITEC convolutional code. In this case, there is little if any performance improvement over the uniformly coded case. The graph shows that the distortion provided by any of the choices of m is essentially the same for R less than 1.75. Above this rate, the uniform coder provides distinctly better performance. The calculated values used to plot the curves in Figure 4.7 actually show that there is a very slight performance gain for the $m \times m$ square subsets between $R=1$ and $R = 2.5$ bits/pixel. The savings is only on the order of 1%. For the (8,7) DITEC code, the rate of the code (1.14 channel bits/source bit) is so close to unity that selecting specific subsets of coefficients does not result in a significant savings in the number of transmitted bits.

The next subset of DCT coefficients chosen was a triangular shaped set of the lowest frequency terms. The size of the triangle (t) indicates the number of elements on its side (i.e. all coefficients with $u + v \leq t-1$ are included in the set). For the (7,4) Hamming code (Figure 4.8), this subset selection resulted in lower distortion levels over the range of rates from 1.3 to 6.8 bits/pixel. The triangular subset provided comparable performance to the $m \times m$ square subset. As an example, at $R=3$ bits/pixel, the best triangular subset ($m=8$), which contains 36 coefficients, produces a weighted MSE of 0.0062. The performance figure of the 6×6 square subset at this rate is 0.0068. Again, the results for the (8,7) DITEC code (Figure 4.9) show that there is little to be gained from coding any subset smaller than the entire block of DCT coefficients.

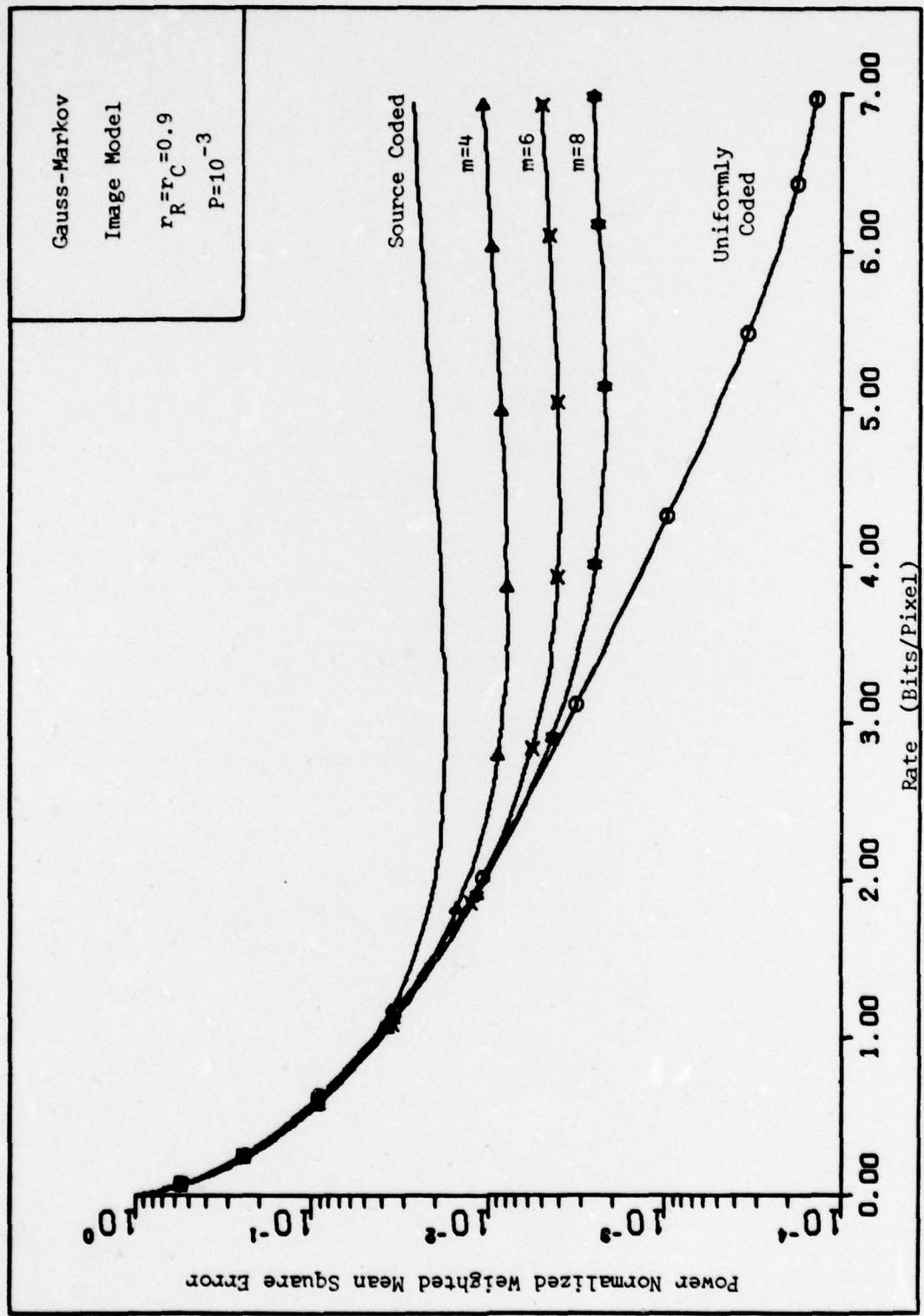


Figure 4.7 Performance of Source/Channel Encoded Imagery Transmission System with an $m \times m$ Square Block of Low Frequency DCT Coefficients (8,7) DITEC Coded

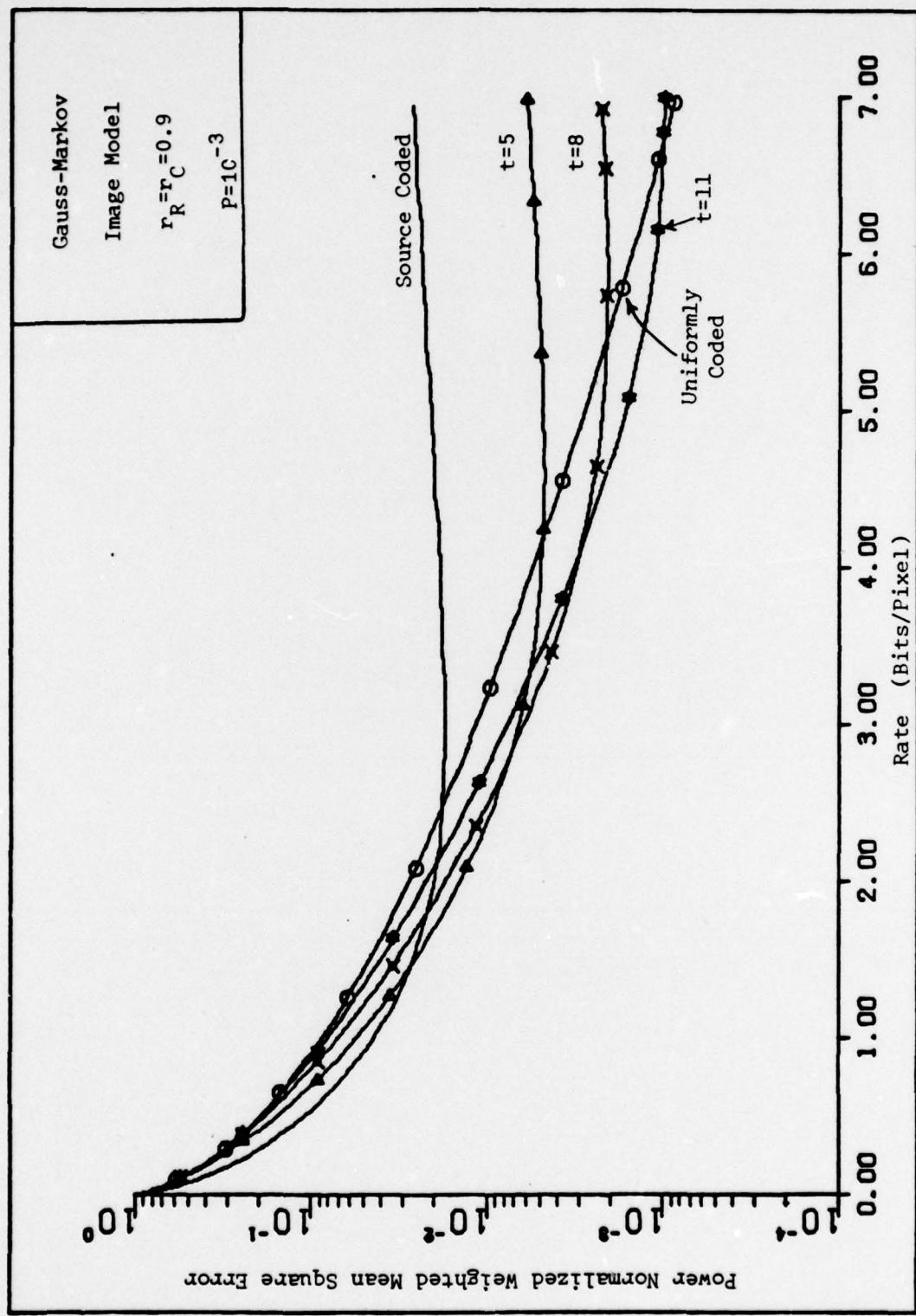


Figure 4.8 Performance of Source/Channel Encoded Imagery Transmission System with a
 $t \times t$ Triangular Block of Low Frequency DCT Coefficients (7,4) Hamming Coded

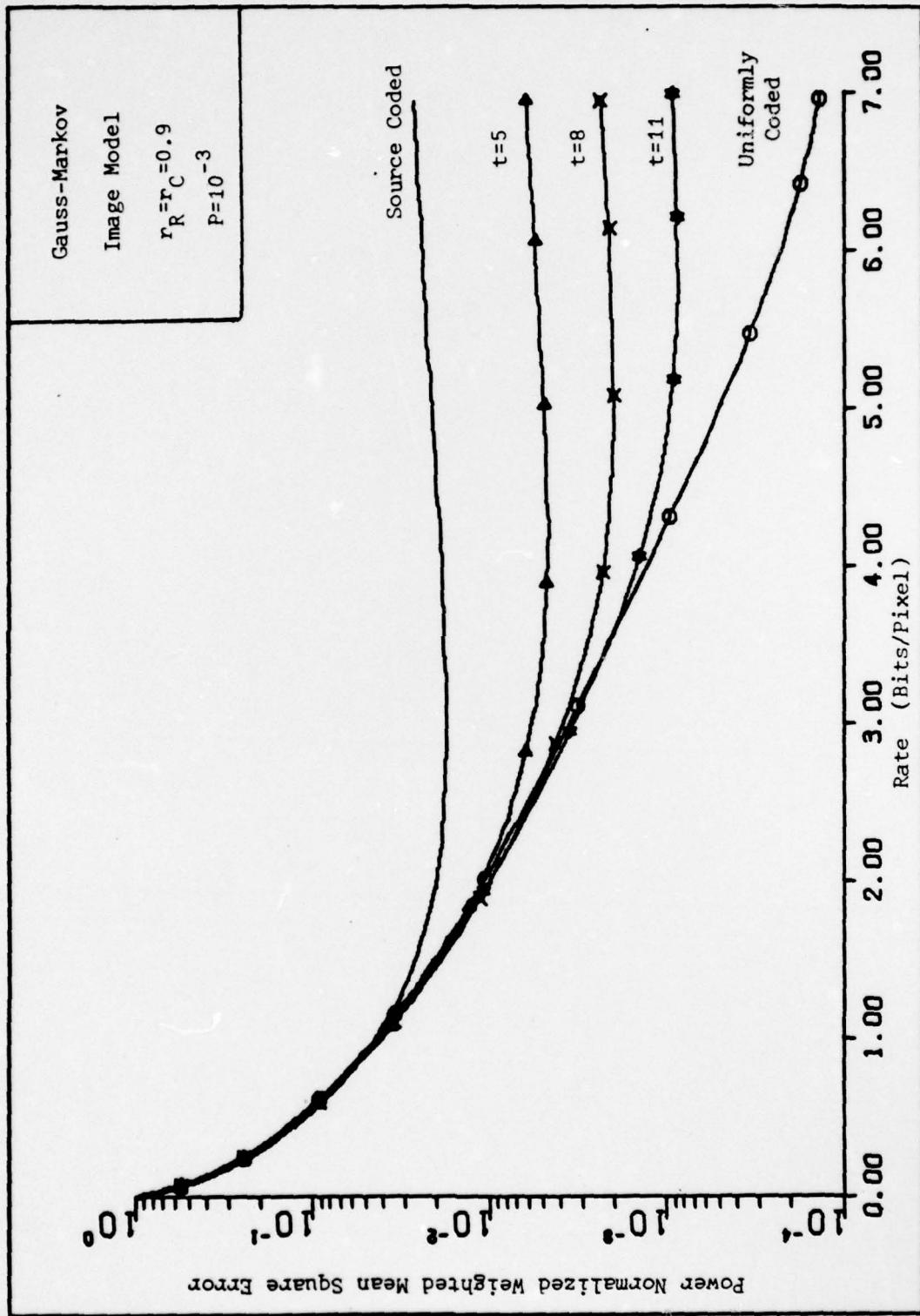


Figure 4.9 Performance of a Source/Channel Encoded Imagery Transmission System with a
 t x t Triangular Block of Low Frequency DCT Coefficients (8,7) DITEC Coded

The final subsets investigated were those which include only the coefficients in the first j rows and j columns. The performance curves for $j = 1, 2$, and 3 are depicted in Figure 4.10 for the $(7,4)$ Hamming code and Figure 4.11 for the $(8,7)$ DITEC code. The Hamming code curves show the typical pattern of performance crossovers among the choices of j . The range of performance improvement extends from 1.5 to 7.0 bits/pixel. At $R=3$ the best distortion level is achieved by the $j=2$ subset and has a weighted MSE of 0.0069. The curves for the $(8,7)$ code again support the conclusion that coding all of the coefficients provides nearly the best performance for all rates.

In summary, channel coding specific subsets of the DCT coefficients can result in improved performance in comparison to the uniform application of the channel code. However, as the rate of the code approaches unity, the magnitude of the performance gain decreases quickly. Also, no one particular subset of the three investigated was found to provide a significant performance advantage as had been anticipated.

Application of Error Correcting Codes to the Most Significant Bits of the Quantized DCT Coefficients. The final method of applying channel codes to the source encoder output is to protect only the s most significant bits (MSB) of each quantizer output word. The equation used to calculate the rate for this method is:

$$R = \frac{1}{M^2} \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} n_r(u,v) \quad (4.9)$$

where

$$n_r(u,v) = \begin{cases} N/K n(u,v) & , \quad n(u,v) \leq s \\ n(u,v)+(N-K)s & , \quad n(u,v) > s \end{cases}$$

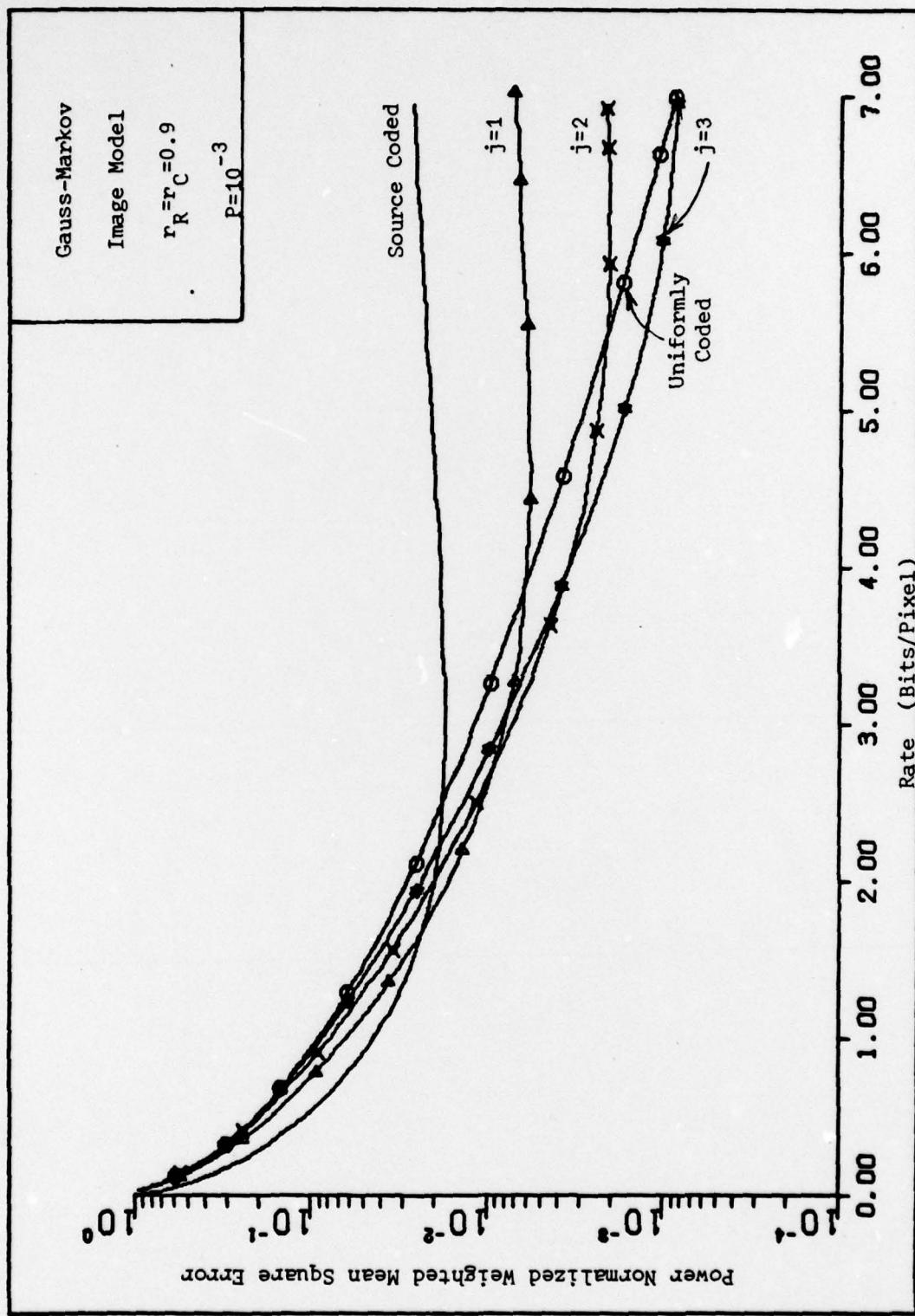


Figure 4.10 Performance of Source/Channel Encoded Imagery Transmission System with j Rows and j Columns of Low Frequency DCT Coefficients (7,4) Hamming Coded

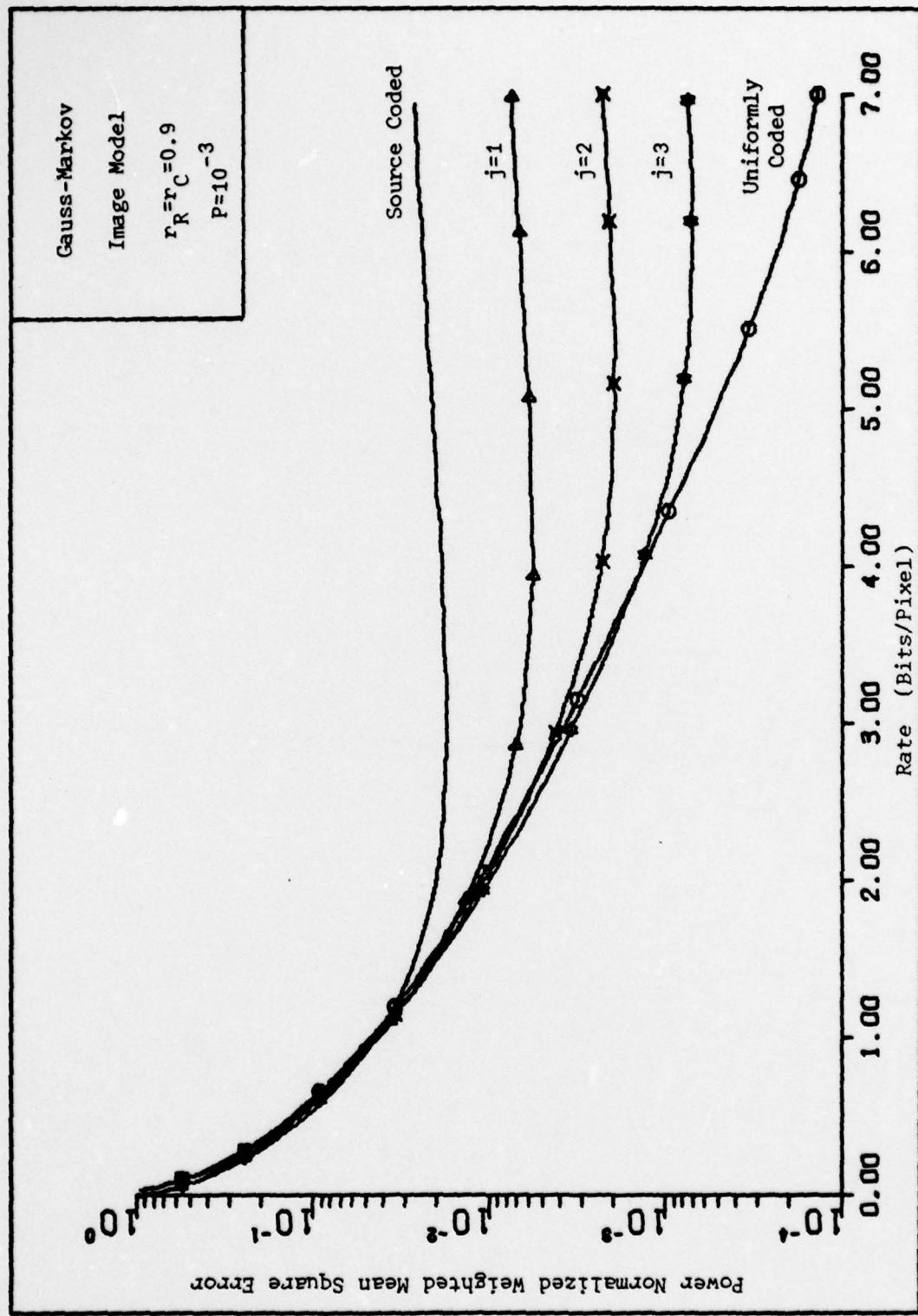


Figure 4.11 Performance of Source/Channel Encoded Imagery Transmission System with j Rows and j Columns of Low Frequency DCT Coefficients (8,7) DITEC Coded

The distortion equation for this application of channel codes was developed in Chapter III with the final result given in equation (3.66). This equation is reproduced here for convenience

$$D_T = \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} \frac{A^2(u,v)}{4} \sigma^2(u,v) \left\{ Q^2[n(u,v)] + d^2(u,v) \left(\sum_{i=0}^{n-1} P_i (1-P_i) 2^{2i} + \sum_{l=0}^{L(u,v)-1} \left[\sum_{i=0}^{n(u,v)-1} \sum_{j=0}^{n(u,v)-1} (1-2l_i) P_i (1-2l_j) P_j 2^{i+j} \right] P(l) \right) \right\} \quad (4.10)$$

where

$$P_i = \begin{cases} P_e \text{ of the } (N,K) \text{ code for the } s \text{ MSB's} \\ 10^{-3} \text{ for all other bits} \end{cases}$$

The term in parentheses which multiples $d^2(u,v)$ can be calculated as a function of n and tabulated once the P_i assignments are made.

As previously explained, only the (3,1) repetition code and the (2,1) Mariner code were considered to be applicable to the task of coding individual bits. The calculated performance curves for the two codes with $s = 1, 2$ and 3 are presented in Figures 4.12 and 4.13, respectively.

For both codes, the main conclusion that can be drawn from these graphs is that protecting even one bit of each quantizer word significantly reduces the weighted MSE of the received image. In fact, for the (3,1) code, protecting only the most significant bit results in better performance over the range from 0 to 7 bits/pixel than coding all bits. For the (2,1) code, the $s=1$ distortion levels are lower than the uniformly coded values for

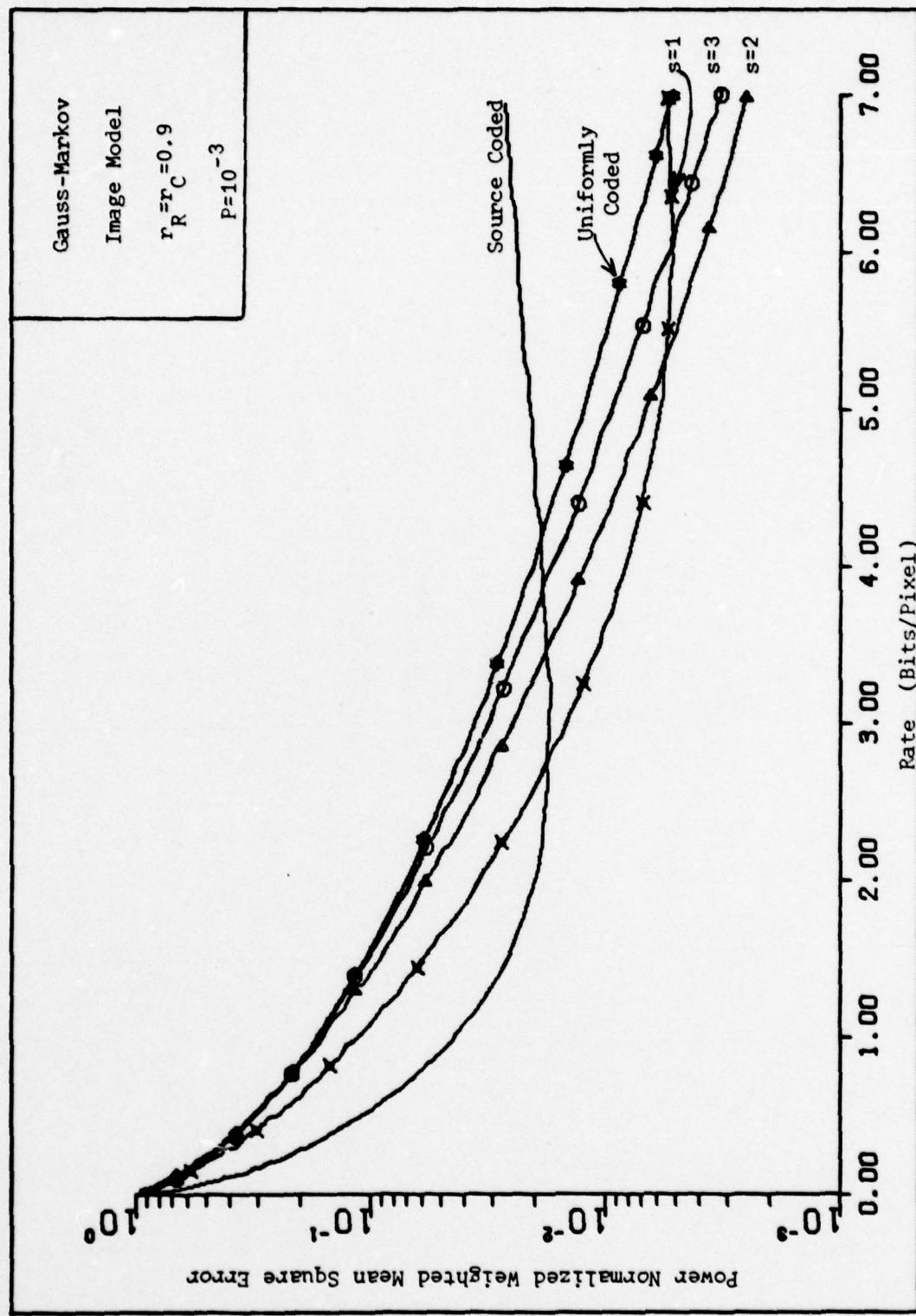


Figure 4.12 Performance of Source/Channel Encoded Imagery Transmission System with s of the Most Significant Bits of Each Quantized DCT Coefficient (3,1) Repetition Coded

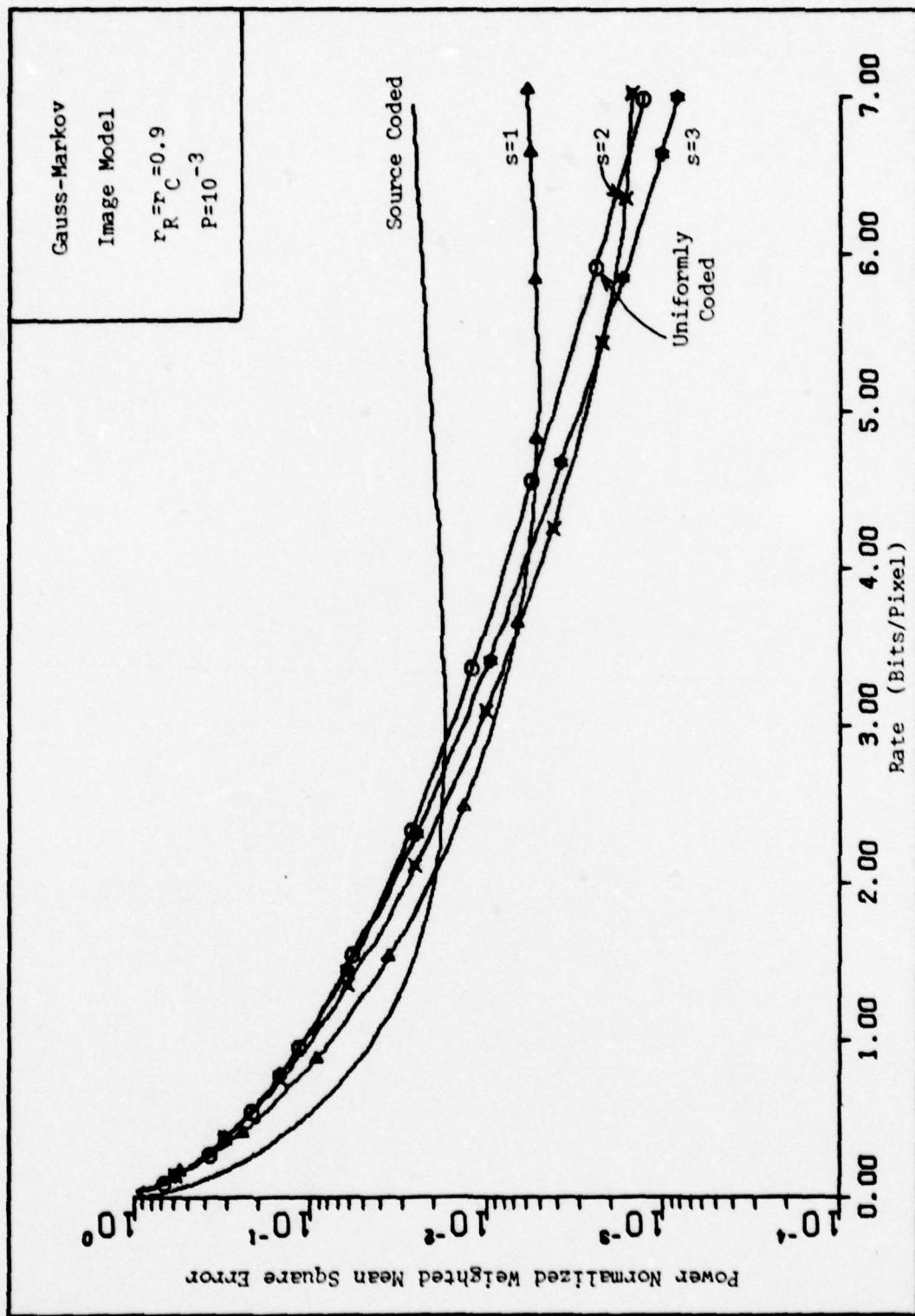


Figure 4.13 Performance of Source/Channel Encoded Imagery Transmission System with s of the Most Significant Bits of Each Quantized DCT Coefficient (2,1) Mariner Coded

all rates from 0 to 4.5 bits/pixel. At a rate of 3 bits/pixel, application of the (3,1) code to all bits provides a higher WMSE than the source encoder alone. Using that same code, on the most significant bit, results in a 20% reduction in the distortion level.

Summary of Source/Channel Coder Performance

In order to tie the results of Chapters II, III and IV together, Figure 4.14 presents the performance of the best overall source/channel encoder in comparison to the performance of the source encoder alone and in comparison to the rate distortion bound for the image source model. The best source/channel coding combination was the (8,7) DITEC convolutional code applied uniformly to all source bits. The performance curve for this combination at $P=10^{-3}$ and the source coder's performance at this bit error rate are shown in Figure 4.14. Also plotted is the source encoder's performance over an error free channel. This curve represents the best performance that can be expected based on the structure of the source encoder.

The graph shows that the application of efficient channel coding procedures can provide performance on even very noisy communications circuits which is very close to the maximum performance of the source encoder. Comparing the source/channel coder's performance to the rate distortion bound, shows that there is still room for performance improvement. These performance gains will have to come through improvements to the source encoder. One change which might lead to reduced distortion levels would be to optimally quantize the DCT coefficients rather than uniformly quantize them. However, no matter what source encoding procedures are employed, the potential gains to be achieved are small in comparison to the improvement in distortion levels that channel coding can provide.

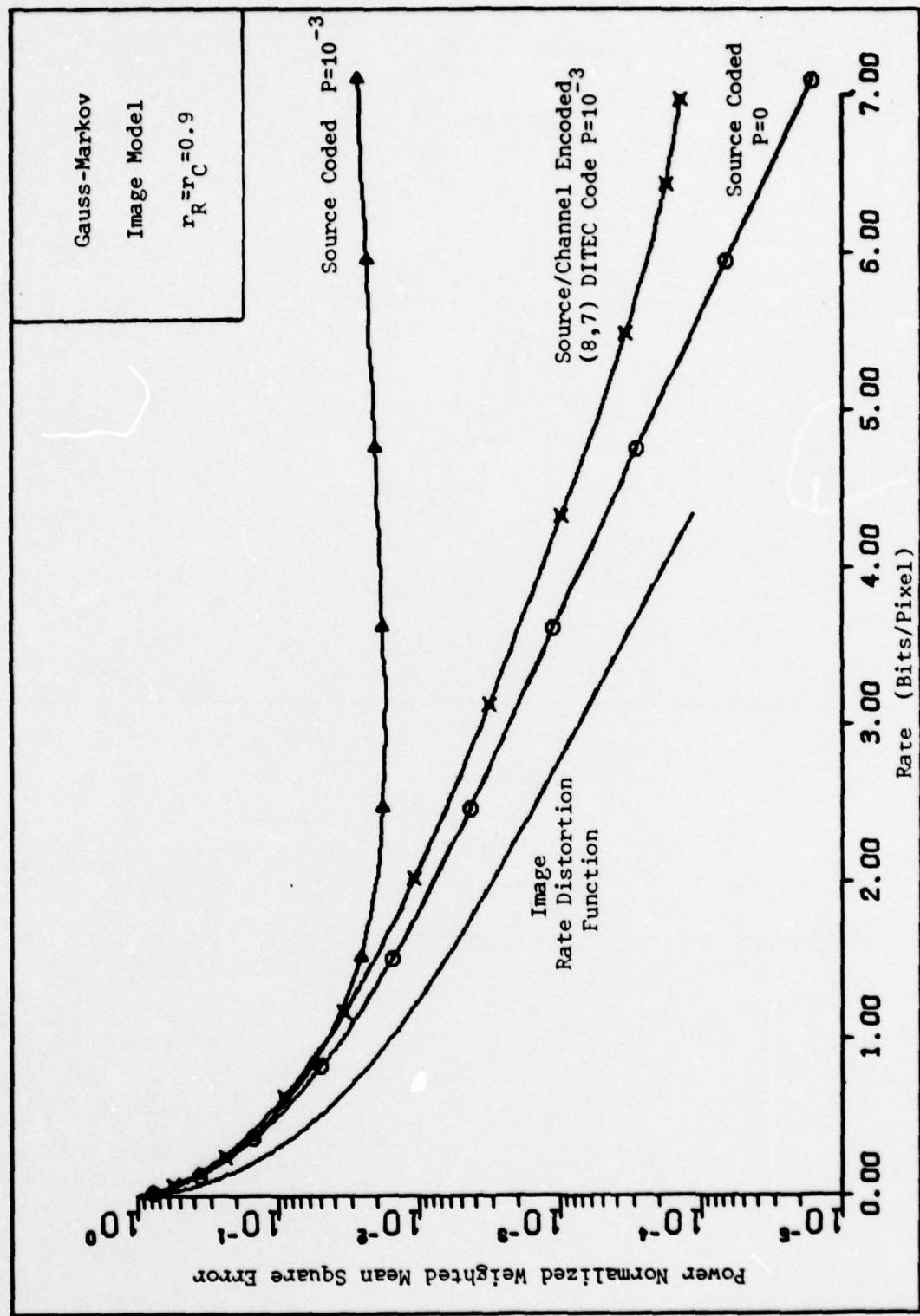


Figure 4.14 Comparison of Source Encoder and Source/Channel Encoder Performance to the Rate Distortion Bound (WMSE)

Finally, Figure 4.15 shows the same four curves as Figure 4.14, except in this case the distortion measure is mean square error. The two figures are nearly identical. Careful examination reveals the MSE curves always lie below the corresponding WMSE curves. This is because the power normalized weighting function is used in the WMSE curves. In terms of comparing the relative performance of the source encoder to the source/channel encoder, it appears that either distortion measure would be suitable.

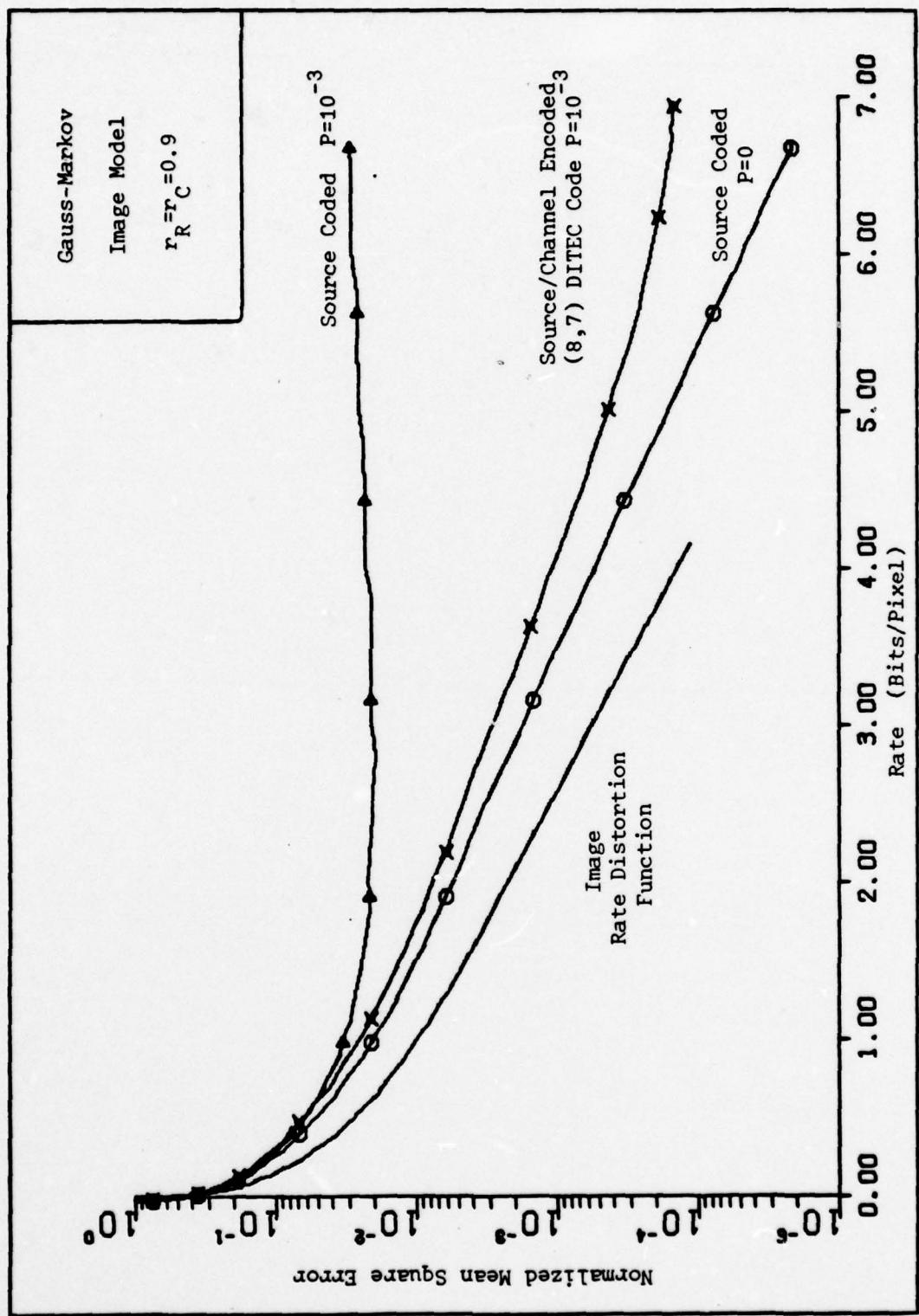


Figure 4.15 Comparison of Source Encoder and Source/Channel Encoder Performance to the Rate Distortion Bound (MSE)

V. Image Transmission Simulation

All of the results presented in Chapters II, III and IV are based on the Gauss-Markov image model. In order to verify these results for real images, a simulated imagery transmission system was established. The simulated system was used to process example images through a discrete cosine transform/block quantization source encoder. The source data bits were then "transmitted" through a simulated communication channel. The channel's bit error probability was varied to simulate the effect of adding error correcting codes to the data stream. The cosine transform coefficients were reconstructed from the received data stream and inverse transformed to produce the output image.

This chapter describes the hardware and software used in the simulation. It also presents examples of the received images and graphs of the measured distortion in the simulated transmissions.

Simulation Hardware

The simulated imagery transmission system was implemented on computer and image processing facilities of the Air Force Avionics Laboratory (AFAL). The Laboratory has a PDP-11 based Image Processing System which can digitize hard copy images, perform some limited statistical manipulations on the image data, display the results on a television monitor and record images on film.

The Image Processing System has a DICOMED image scanner and recorder. The DICOMED scanner can digitize a 57 x 57 mm area of an image. It provides several different scanning resolutions. The highest resolution setting produces 2048 samples across a 57 mm scan line and 2048 scan lines per image.

The image data can be quantized to either 6 or 8 bits per pixel which correspond to 64 and 256 gray levels, respectively.

The DICOMED recorder can produce a 54 x 54 mm hard copy image. The recorder always prints 2048 by 2048 pixels over the area. The recorder does have the capability to repeat pixels in square arrays to reduce the effective resolution of the recorded image. The lowest resolution setting converts each incoming pixel into a 4 x 4 square array of pixels. This yields an effective resolution of 512 lines over the 54 mm square image area.

The Image Processing System was used primarily to digitize the original images and to record the images produced by the simulated transmission system. As described in Chapter II, the desired resolution for the imagery is 60 lines per inch. For a 54 mm image, this equates to 128 samples per line. This is a very convenient number since it is an exact multiple of 16, the selected block size of the DCT transform. In order to produce this resolution on the DICOMED recorder, each pixel must be repeated 16 times. The Image Processing System has a utility function which allows an image to be magnified by repeating pixels and lines. By using this function to expand the 128 x 128 image by a factor of 4 and then using the low resolution setting of the image recorder, the required 16 times magnification can be achieved.

The input images to the transmission simulation were prepared by scanning the original hard copy pictures at the maximum resolution of the DICOMED scanner, 2048 pixels/scan line. The 128 x 128 pixel images needed for the simulation were produced from the scanned data by breaking the samples into 16 x 16 square blocks. The 256 samples in each block then were averaged to produce a single 8 bit pixel value. This procedure was

followed to eliminate aliasing effects in the digitized images. The averaging operation produces a $\sin x/x$ filter response in the frequency domain. This is not a particularly good low pass filter, due to the significant side lobe structure which it allows. However, the side lobes are located at spatial frequencies where the eye has little sensitivity. Therefore, the quality of the imagery produced by the averaging operation is adequate for the purposes of the study.

After the images were scanned and averaged, the resultant 128 x 128 sample images were transferred by magnetic tape from the Image Processing System to the AFAL DEC-10 computer. The DEC-10 was used to simulate the source/channel encoder and to "transmit" the sample images. The "received" images then were transferred back to the Image Processing System by magnetic tape to be displayed and recorded. The DEC-10 was used for the actual simulation, since it provides many more utility functions than the PDP-11, and can easily handle the large data files and arrays needed to store and process imagery.

Simulation Software

Three FORTRAN programs were written and used to perform the image transmission simulation. The first program performs the discrete cosine transform on the input image. The second calculates the block quantizer bit assignment matrix. The third program takes the results of the first two, quantizes the DCT coefficients, transmits the quantized values through a simulated communication channel, and then inverse transforms the received data to produce the output image. Each of these three programs is discussed below.

Transform Program. Since the primary emphasis of the simulation is to determine the effect of adding channel encoding to the transmitted data

stream, little variation was allowed in the input images. Only three images were used in the simulation, and one image in particular was used to demonstrate the impact of channel encoding. While the number of images was limited, the number of simulated transmission was quite large. Each image was repeatedly transmitted while the parameters of the simulation were varied. A program was written to eliminate the need to perform the discrete cosine transform each time an image was transmitted. The program performs the transform once for each image and stores the resultant DCT coefficients.

The Transform program performs its required operation in four separate steps. First, it accepts the input image and calculates a sample mean, variance and row and column correlation coefficients for it. This segment of the program also subtracts the sample mean from each pixel. Then, using the sample variance and correlation coefficients, estimates of the variances of the DCT coefficients, $\sigma^2(u,v)$, are calculated in accordance with equation (3.9). Next, the 128 by 128 image values are divided into 64, 16 x 16 blocks, and the two dimensional discrete cosine transform performed on each block. The 2-D transform is accomplished by doing a 1-D transform along the rows of a block and then performing a second 1-D transform across the columns of the block. Both 1-D transforms are calculated using the fast algorithm described in Chapter III. Finally, each of the DCT coefficients is divided by its estimated standard deviation, as required by equation (3.25). The normalized DCT coefficients, along with the sample mean, variance and correlation coefficients, are stored in a file for use by the other two programs.

Bit Matrix Program. The second simulation program calculates the number of bits to assign to each DCT coefficient. The program is based on

the bit assignment procedure, described in Chapter III, which incorporates equation (3.24). That is, the number of bits assigned to each coefficient depends on the weighted variance of the coefficient, $A^2(u,v) \sigma^2(u,v)$, and a pre-established threshold, Q . The variance and correlation coefficients determined in the Transform program are used to calculate estimated weighted variances using equations (3.8) and (2.13). Then, given a Q value, the program calculates a bit assignment matrix and displays the average rate in bits per pixel. It also displays distortion values which are calculated assuming the image source is a Gauss-Markov process. Q values then can be entered interactively to set either the distortion or average rate to a desired level.

In the simulation, Q was adjusted to give exact integer values for the average rate. By fixing the rate to integer values, comparison of the distortion levels for various combinations of source/channel codes is greatly facilitated. Also note that calculation of the bit assignment matrix, in this manner, eliminates any requirement to manually "adjust" the bit assignments. The algorithm automatically assigns the bits in a manner which produces the lowest distortion. Once the bit assignment matrix for the desired rate is found, it is written into a temporary file for use by the Transmit program described below.

Several modifications were made to the Bit Matrix program to incorporate channel encoding. Simple logical statements can be added to indicate which DCT coefficients or which bits within a quantizer word are to be protected. Then, the total number of source bits to be channel encoded is determined and that sum multiplied by the rate of the channel code. This product is added to the number of bits not protected. The resultant sum is used to calculate the average rate with channel encoding.

It must be noted that the calculated rates do not include any contribution for overhead or synchronization bits. These would be necessary in a practical image transmission system. There is no reason to expect that the number of overhead bits would change significantly with variations in the source/channel encoder. Therefore, in comparing the relative performance of different source/channel encoder combinations, omission of the overhead bits has no effect. In absolute terms, the calculated transmission rates would be greater for a practical system. Data from operational image transmission systems indicates that the overhead bits increase the average rate by 10% or less (Ref 40:48).

Transmit Program. The third and final program, the Transmit program, takes the normalized DCT coefficients calculated by the Transform program, quantizes them to the number of levels dictated by the bit assignment matrix and then "transmits" them through a simulated communication channel. It reconstructs the DCT coefficients from the received data and rescales them by multiplying each DCT coefficient by its estimated variance. These DCT values are then inverse transformed. The final steps in the program add the sample mean to each of the received pixel values and store them in an output file. The program is written under the assumption that the receiver knows the mean, variance and correlation coefficients of the incoming image information. In practice, this could easily be accommodated by transmitting this information, with appropriate error protection, as part of the image data.

The only parts of the Transmit program which need to be described in detail are the quantizer and the simulated communication channel. The quantizer is a straightforward implementation of the uniform quantizer depicted in Figure 3.7. Each of the normalized DCT coefficients is divided

by the quantization interval width, d , which is dictated by the number of bits assigned to the coefficient, n (Ref Table III.4). The result is truncated to an integer and scaled upward by adding $L/2$ ($L=2^n$) to give the quantizer level. The result is checked to determine if it is less than 0 or greater than $L-1$. That is, a test is made to determine if the value of the DCT coefficient falls outside the input range of the quantizer. If so, the coefficient is assigned a quantization level of 0 or $L-1$, appropriately. The quantization level expressed as an integer then is fed into the simulated communication channel. The received value, also expressed as an integer, is scaled down by subtracting $(L-1)/2$ and then multiplied by d to give the normalized DCT output value.

The simulated communication channel routine is depicted in flow chart form in Figure 5.1. The simulated link is a discrete, memoryless binary symmetric channel. The probability of an error occurring in a bit (P_i) is dependent on which bit is being processed. By changing certain logical statements within the Transmit program, these bit error probabilities can be controlled to simulate the addition of error correction codes to all transmitted DCT coefficients, specific subsets of DCT coefficients or to individual bits within the quantizer output words.

The communication channel routine operates in the following manner. The quantization level, expressed as an integer (L_Q), along with the number of bits assigned to the corresponding DCT coefficient, n , are input to the routine. The algorithm determines whether each bit in the n bit input value is a 1 or a 0. Starting with the most significant bit, the routine iteratively divides L_Q by 2^{n-i} ($i=1, 2, \dots, n$) and tests the quotient to determine if it is greater or less than 1. If it is greater than or equal to 1, the corresponding bit is a 1, so 2^{n-i} is subtracted

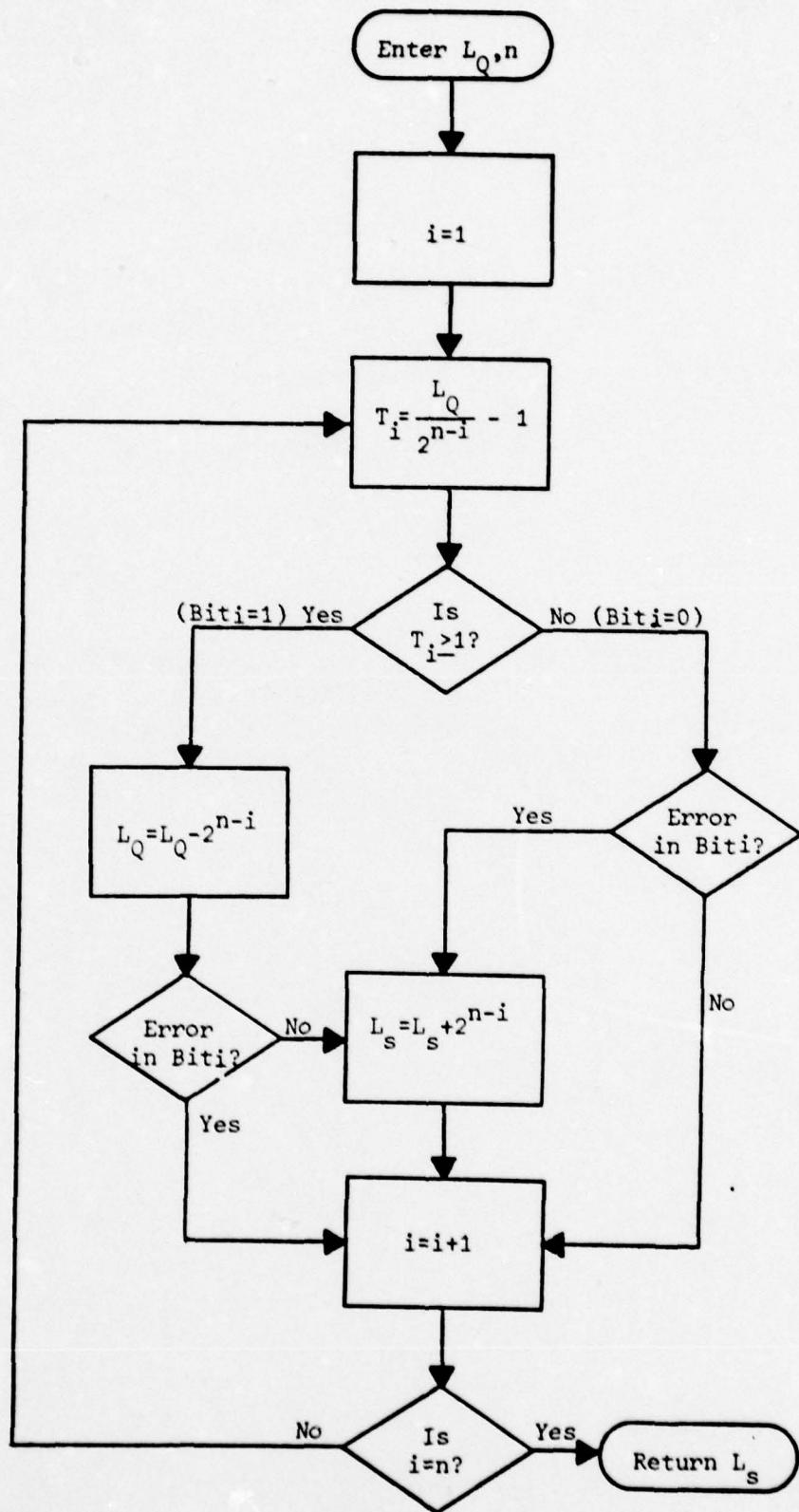


Figure 5.1 Flow Chart of Simulated Communication Channel

from L_Q and the process continued. If the quotient is less than 1, the corresponding bit is a 0, so L_Q is left unchanged. The routine then determines if an error occurs in that bit by obtaining a random number between 0 and 1 from the standard FORTRAN random number generator. This value is compared to P_i . If it is less than P_i , the bit is complemented. Finally, the value of the transmitted bit multiplied by 2^{n-i} is added to a running total, which is maintained throughout the n iterations of the routine. This sum is output as the received quantization level at the completion of the routine. The procedure was specifically designed to operate on the most significant bits first to simplify the simulation of channel encoding specific bits in the quantizer word. For example, all that needs to be done to simulate the protection of the most significant bit is to change P_1 , to the effective error rate, P_e , of the channel code.

The Transmit program also calculates the weighted and unweighted mean square error in the received image. As in the theoretical performance calculations, the weighted mean square error is determined in the frequency domain by comparing the DCT values before and after transmission (equation 3.19). The normalized mean square error is calculated using equation (2.7). The weighted mean square error is already normalized by the choice of the weighting function $A(u,v)$. In most cases, the power normalized version $A_p(u,v)$, given by equation (2.24), was used. However, for some results, the amplitude normalized weighting function $A_n(u,v)$ defined by equation (2.25) was used. The specific version of the weighting function will be defined for each set of results presented below.

Simulation Results

A total of more than 200 simulated image transmissions were performed.

Many of these were repetitions of each other. The repetitions were required to determine an average level of distortion for a specific combination of source/channel encoding procedures, average transmission rate and channel error rate. The repetition of transmissions became particularly important when the error rate of the communication channel was greater than 10^{-4} . At these error rates, measured distortion in the received image varied considerably from one transmission to the next. Therefore, an image was often transmitted as many as five times under the same set of conditions to obtain an average distortion figure.

The results of these simulated transmissions are presented in graphical form and in sets of example images throughout the remainder of this chapter. In all cases, the distortion figures given are the averages over all transmissions for the particular set of transmission conditions. The specific images displayed in the report are the ones which had measured distortion levels close to the average.

A second point to be made concerning the images is that they are intended to be viewed at a distance of 30 inches. If they are analyzed at closer range, a patchwork structure is clearly visible. This is caused by the repetition of each of the pixels in a 16×16 pattern to produce the appropriate resolution. At the proper viewing distance, this structure is not evident.

Source Encoder Performance. An initial series of image transmissions were conducted to establish the performance of the source encoder alone. Figures 5.2 through 5.7 present the results of varying the average transmission rate (in bits/pixel) with no channel errors. That is, these figures give the maximum distortion performance achieved by the simulated source encoder. Also shown in the graphs of Figures 5.2 through 5.4 are the theoretical

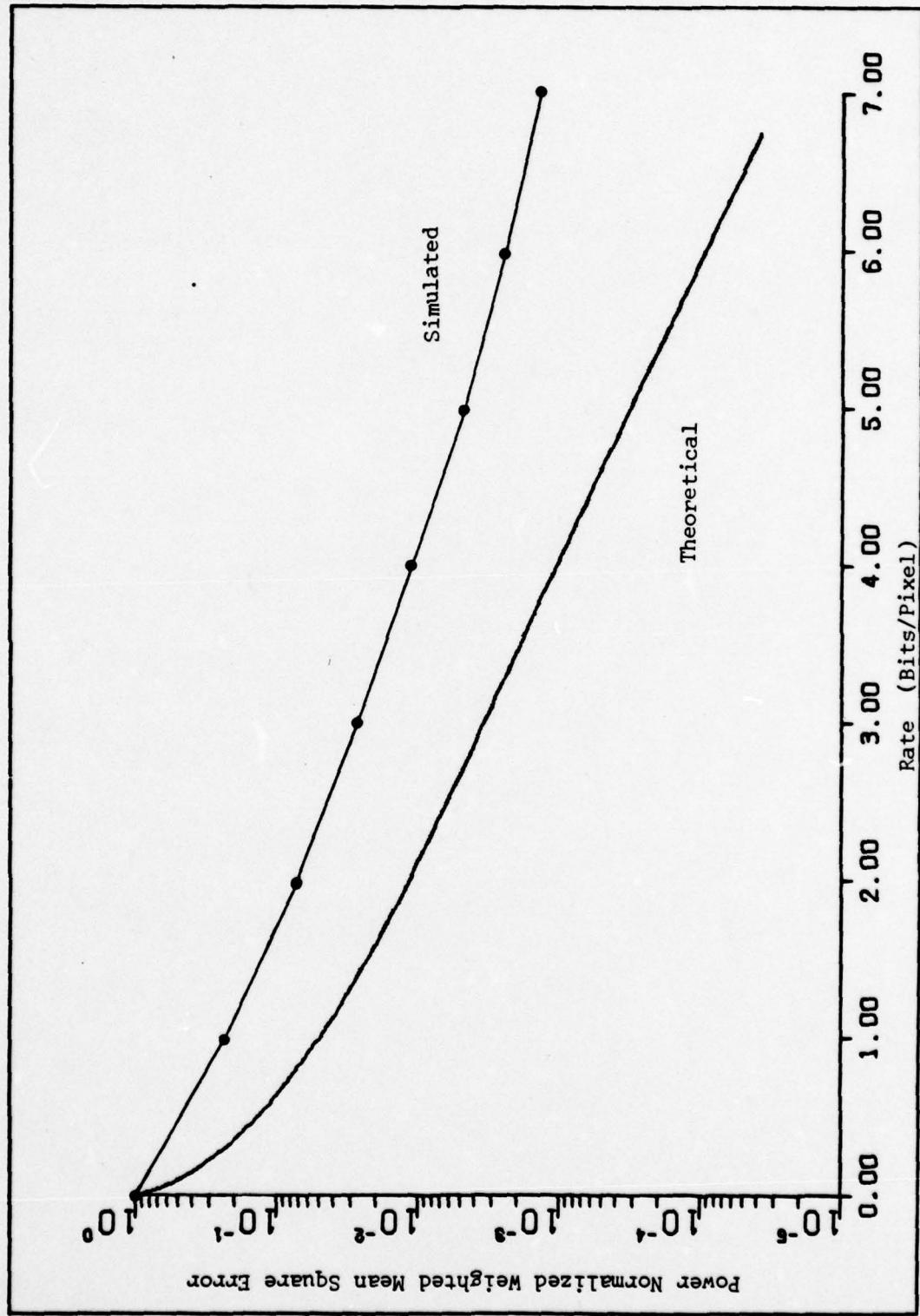


Figure 5.2 Comparison of Theoretical and Simulated Source Encoder Performance for the Truck Image ($r_R = 0.837$; $r_C = 0.858$; $P = 0$)

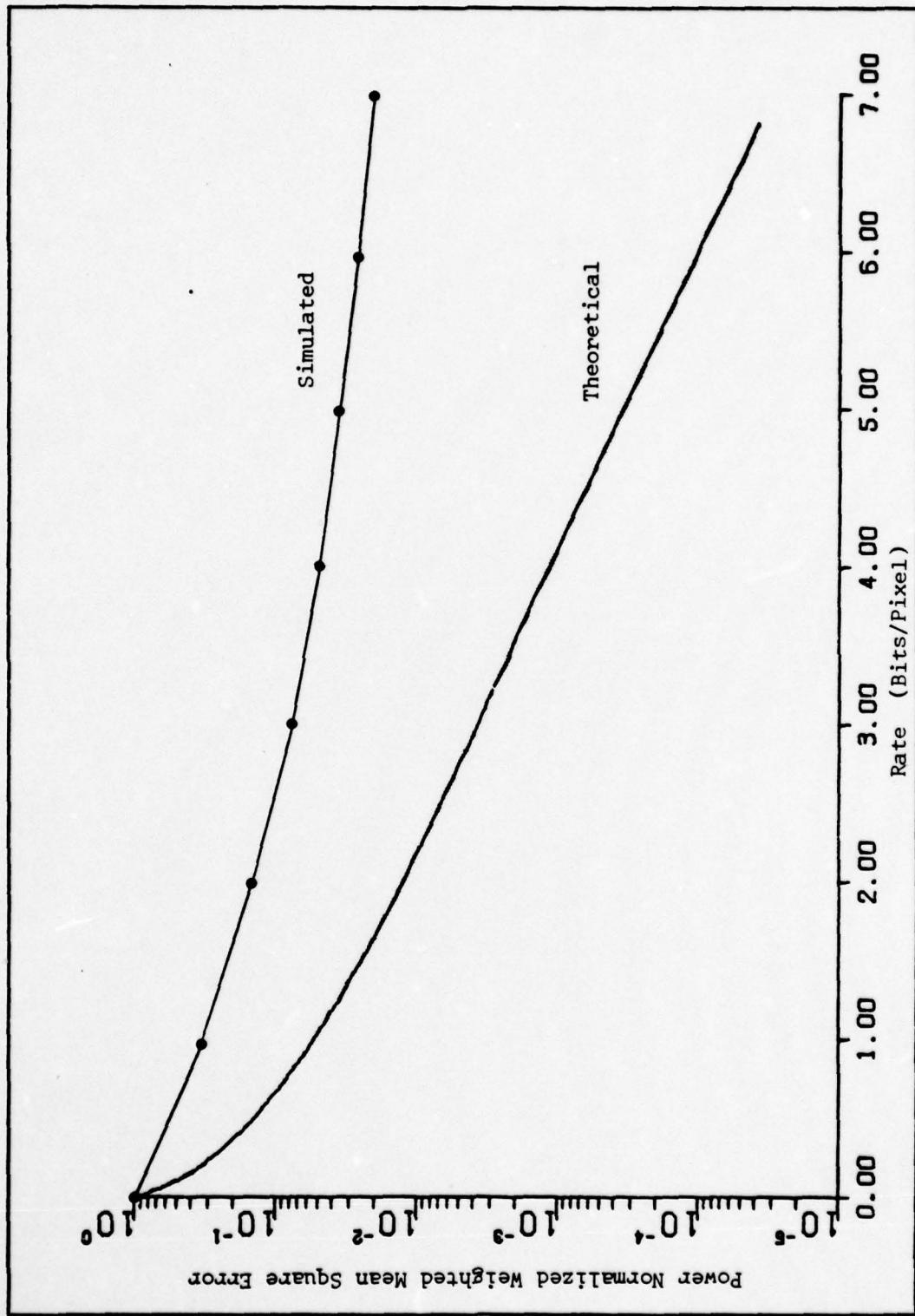


Figure 5.3 Comparison of Simulated and Theoretical Source Encoder Performance for SMSITE Image ($r_R=0.845$, $r_C=0.869$, $P=0$)

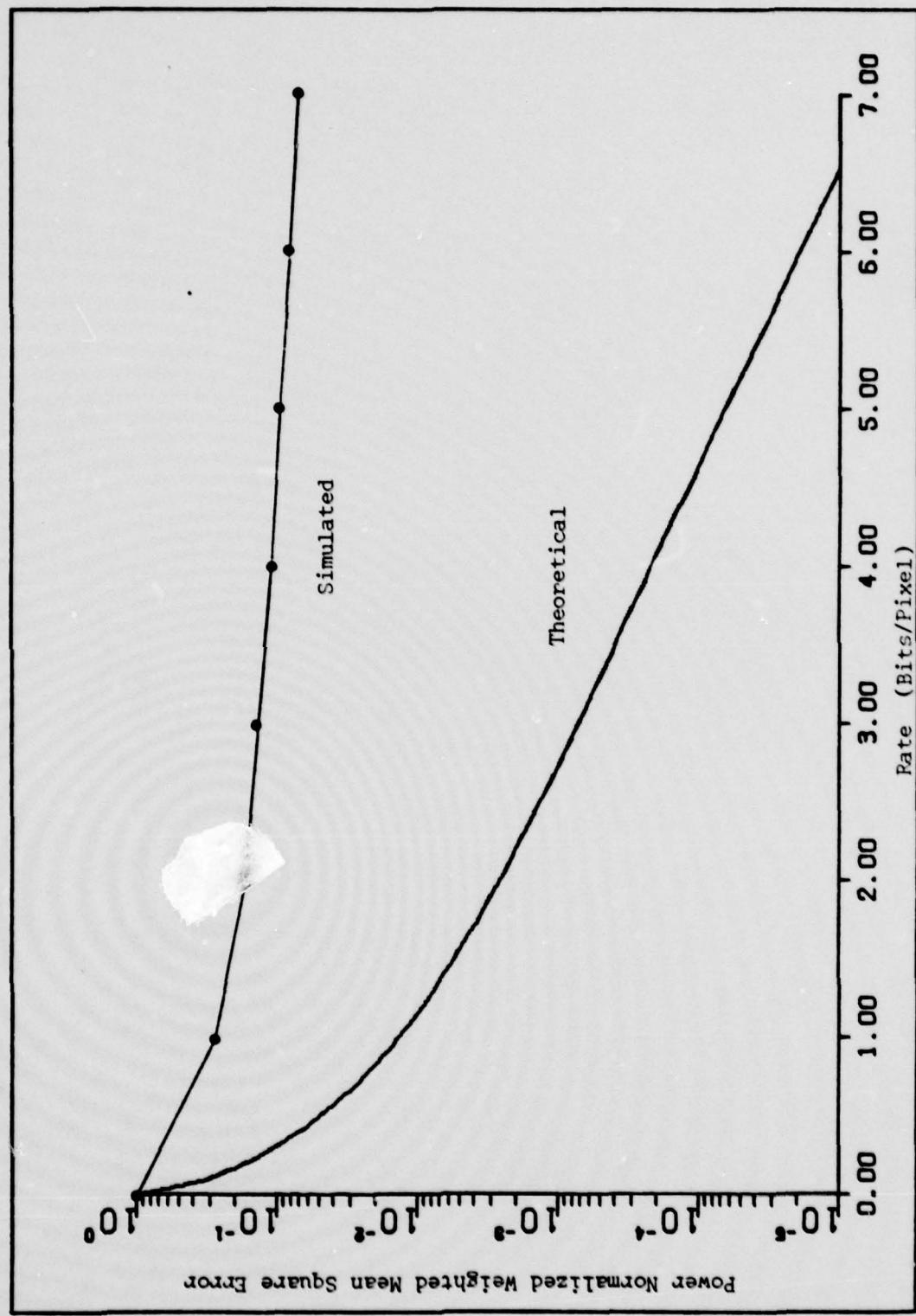


Figure 5.4 Comparison of Theoretical and Simulated Source Encoder Performance for the Bomber Image ($r_p = 0.971$; $r_c = 0.967$; $P = 0$)

performance curves for the Gauss-Markov image model with identical correlation coefficients.

The graphs show that the simulated transmissions have much greater levels of distortion than predicted by the theoretical calculations. The difference is normally an order of magnitude or greater. Some analysis of the simulation was conducted to determine a reason for this large discrepancy.

Careful analysis of the quantizer in operation revealed one potential source for the difference between measured and predicted performance. Observation of the integer values associated with the quantizer showed that many of the normalized DCT values fall outside the input range of the uniform quantizer. This was particularly true when only 1 or 2 bits were assigned to a coefficient. The quantizer automatically assigns either the 0 or L-1 level to these "over-range" values. This causes a large change in the value of the corresponding received DCT coefficient. The contribution of such over-range values to the theoretical distortion is quite small since their probability of occurrence is small. As an example, for $n=2$, the uniform quantizer has an input range of approximately ± 2 . For a 0 mean, unit variance Gaussian random variable, the probability of getting a sample outside this range is only 4.6%. For $n=1$, this probability is approximately 11%. In actual operation, some 15 to 20% of the DCT values were outside the quantizer range on the average.

A probable reason for the large number of over-range values is that the variance used to normalize the DCT coefficients is an estimate. This estimate is calculated under the assumption that the image is Gauss-Markov with variance and correlation coefficients equal to the sample statistics. If the actual measured variance of the individual DCT terms was used as the

normalizing constant, the performance might be improved.

The large distortion caused by the quantizer tends to overshadow the performance gains achievable with channel encoding. In fact, it was necessary to use one image (the Truck image) exclusively in the source/channel encoding simulations in order to show the performance gain. The source encoder's performance is the best for this image.

Figure 5.5 shows the Truck image at different source encoding rates. The DCT/block quantization scheme produces good quality images down to 3 bits/pixel. There are some artifacts of the source encoder (the array of light pixels at the boundaries of the 16 x 16 pixel blocks) in the 4 and 5 bit images. This does not detract from the information content of the image, however. At 2 bits/pixel, the image quality is seriously degraded by the source encoder artifacts and by the elimination of high frequency information.

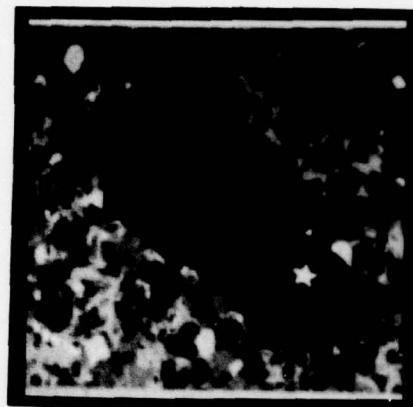
The SMSITE images presented in Figure 5.6 show similar results to the Truck images. The variance of this image is very small in comparison to the other images. Therefore, even small differences in the transmitted and received pixel values result in large weighted mean square error values. The images show little or no degradation down to 3 bits/pixel.

The Bomber images presented in Figure 5.7 show quite different results. Even at 6 bits/pixel the image is seriously degraded by the loss of high frequency information. The correlation coefficients for this image are very high ($r_R = .971$, $r_C = .967$). Therefore, the bit assignment procedure allots most of the bits to the low frequency DCT coefficients and very few to the high ones.

A second series of source encoder transmissions were made to investigate the performance variation under changing communication channel error rates.



Original



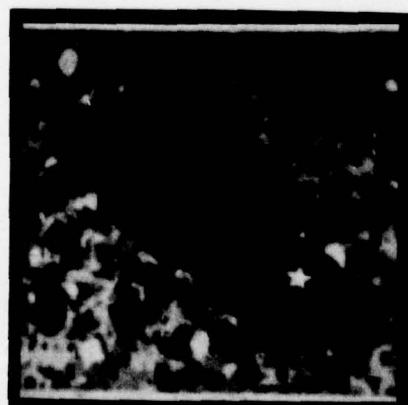
6 Bits .25% .15%



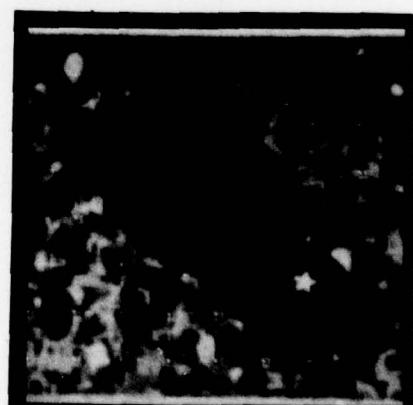
5 Bits .50% .34%



4 Bits 1.14% .84%



3 Bits 2.82% 2.06%



2 Bits 7.36% 6.34%

Figure 5.5 Source Encoded Truck Image Transmitted at P=0
(Distortion figures below each image are power
normalized WMSE and MSE in that order)



Original



6 Bits 2.63% 1.37%



4 Bits 4.92% 2.83%



2 Bits 13.74% 9.00%

Figure 5.6 Source Encoded SMSITE Image Transmitted at P=0
(Distortion Figures below each image are power
normalized WMSE and MSE in that order)



Original



6 Bits 8.17% 1.15%



3 Bits 12.75% 1.91%

Figure 5.7 Source Encoded Bomber Image Transmitted at P=0
(Distortion Figures below each image are power
normalized WMSE and MSE in that order)

The results of these transmissions are shown only in graphical form in Figure 5.8. The performance of the source encoder is fairly good for error rates up to 10^{-4} . For noisier channels, the performance degrades rapidly.

The final set of simulated transmissions for the source encoder are intended to show the effect of using the unweighted MSE criterion, instead of the weighted measure, to make the bit assignments in the block quantizer. Figure 5.9 gives the measured distortion values. Figure 5.10 presents 3 sample images processed with MSE bit assignment matrices.

The results of this series of transmissions are inconclusive. The weighted distortion for the MSE processed images is greater than those processed with weighted bit assignment matrices. However, the measured MSE for the two sets is essentially the same. Comparison of the MSE images of Figure 5.10 to the corresponding WMSE images of Figure 5.4 shows no discernable differences.

The weighted measure for the last set of images discussed employs the amplitude normalized weighting function. Figure 5.9 shows that the weighted distortion is always less than the unweighted error. This is typical of all results obtained in the simulation. It is also in agreement with the theoretical rate distortion curves given in Figure 2.10. Since the peak amplitude of the weighting function is 1, the weighted MSE can not be greater than the MSE. It is possible for them to be equal but this would only occur if the weighting function was very flat across the frequencies of interest. This is not the case for the weighting function used in this report.

A second conclusion that can be drawn from the simulated data is that the power normalized WMSE is always greater than the MSE distortion. What

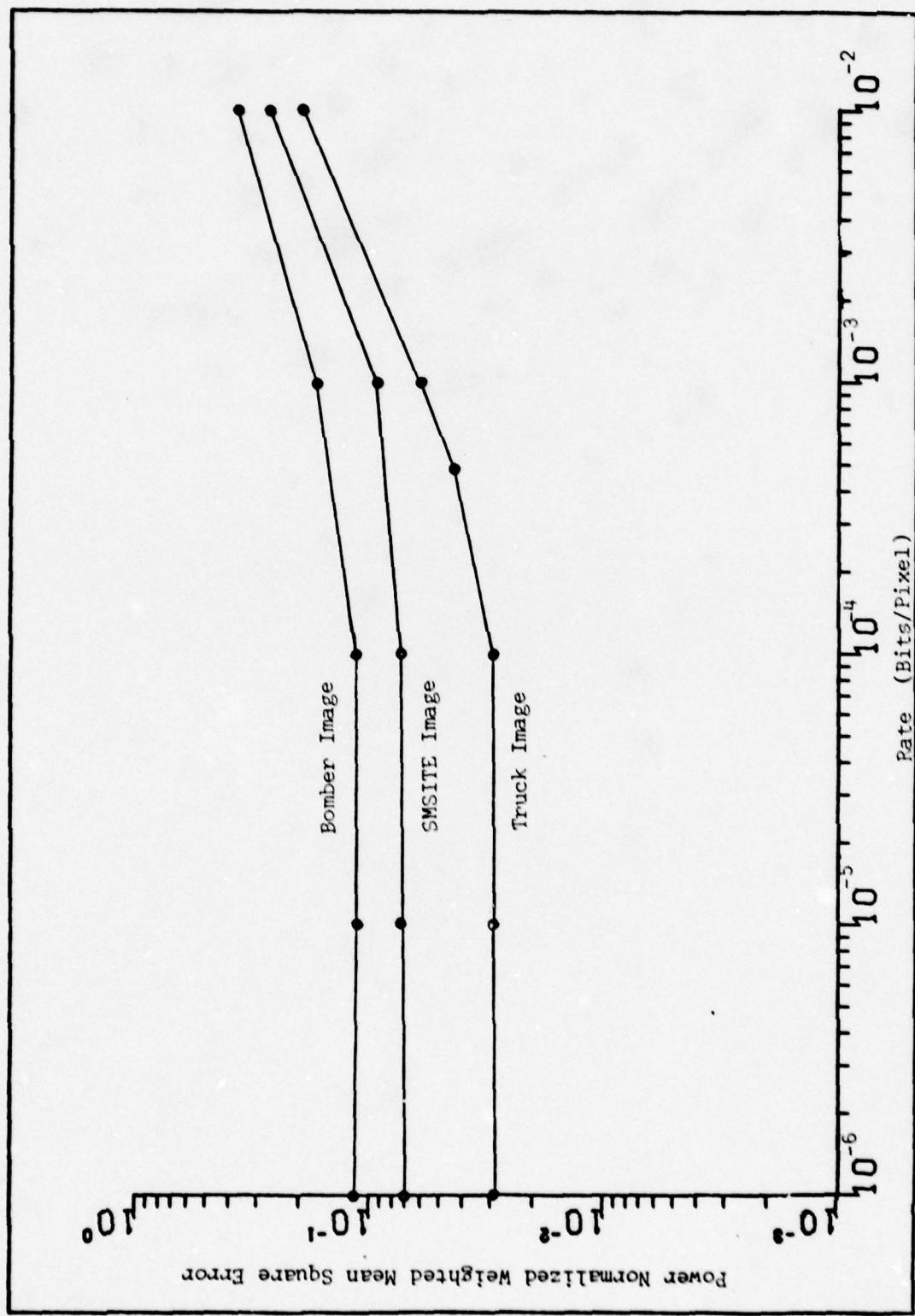


Figure 5.8 Source Encoder Performance Under Varying Communication Channel Error Rates
(Transmission Rate Held Constant at 3 Bits/Pixel)

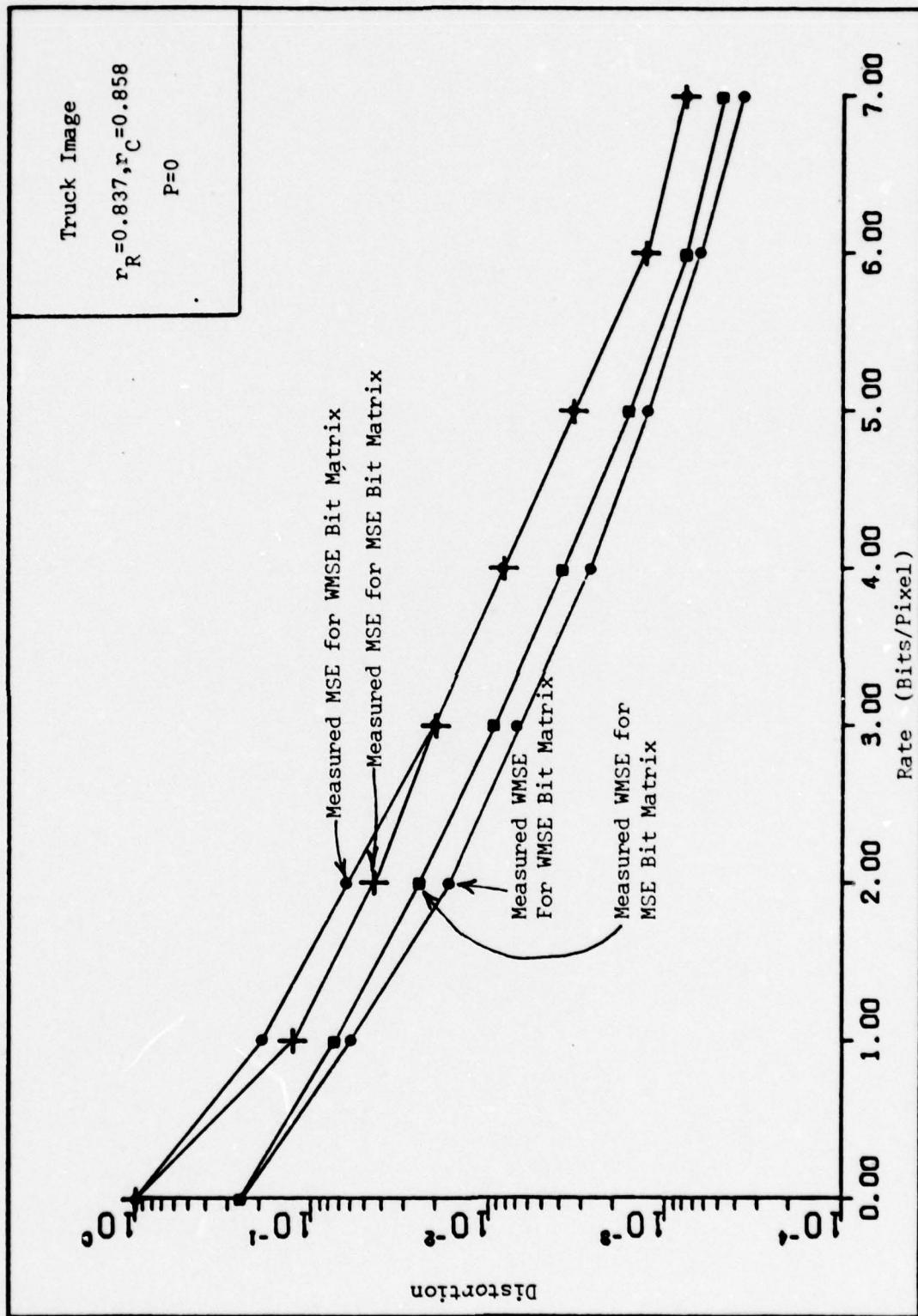


Figure 5.9 Comparison of Simulated Source Coder Performance for WMSE and MSE Bit Assignment Matrices (Amplitude Normalized WMSE Used)



6 Bits .29% .16%



4 Bits 1.49% .85%



2 Bits 9.7% 5.2%

Figure 5.10 Source Encoded Truck Image Transmitted at $P=0$
Using MSE Assigned Bit Matrix
(Distortion Figures below each image are power
normalized WMSE and MSE in that order)

is interesting about this is the power normalized and amplitude normalized weighting functions differ by only a gain constant. The measured contrast threshold curves used as a basis for the weighting function only provide relative sensitivity of the human visual system to spatial frequencies. The absolute scale of the frequency response is unknown. Therefore, the true gain constant associated with the weighting function is also unknown. This makes any comparison of the weighted distortion to the unweighted distortion impossible. There is no way of knowing where the WMSE curve actually lies in relation to the MSE curve. However, either distortion measure is adequate in terms of comparing the relative performance of one source/channel code combination to another.

Source/Channel Encoder Performance. The effect of channel encoding the source encoder's data stream was demonstrated through a number of simulated transmissions. The theoretical results presented in Chapter IV were used to guide the selection of the types of channel codes applied and the procedures for applying them. Since the performance of the source encoder did not degrade significantly until the channel error rate approached 10^{-3} , this error rate was used throughout the source/channel encoding simulation.

In order to establish a baseline for comparison, the source encoder's performance at $P=10^{-3}$ was determined over the range of transmission rates from 1 to 7 bits/pixel. Figure 5.11 shows the measured performance curve. This performance curve exhibits the same characteristics as the theoretical curve of Figure 3.8. Namely, the distortion decreases as the rate increases, up to a certain level. However, beyond that rate the distortion actually worsens as the rate increases. For the simulated data, the minimum distortion is achieved at 5 bits/pixel.

Images a, b and c of Figure 5.12 are samples of the source encoder's

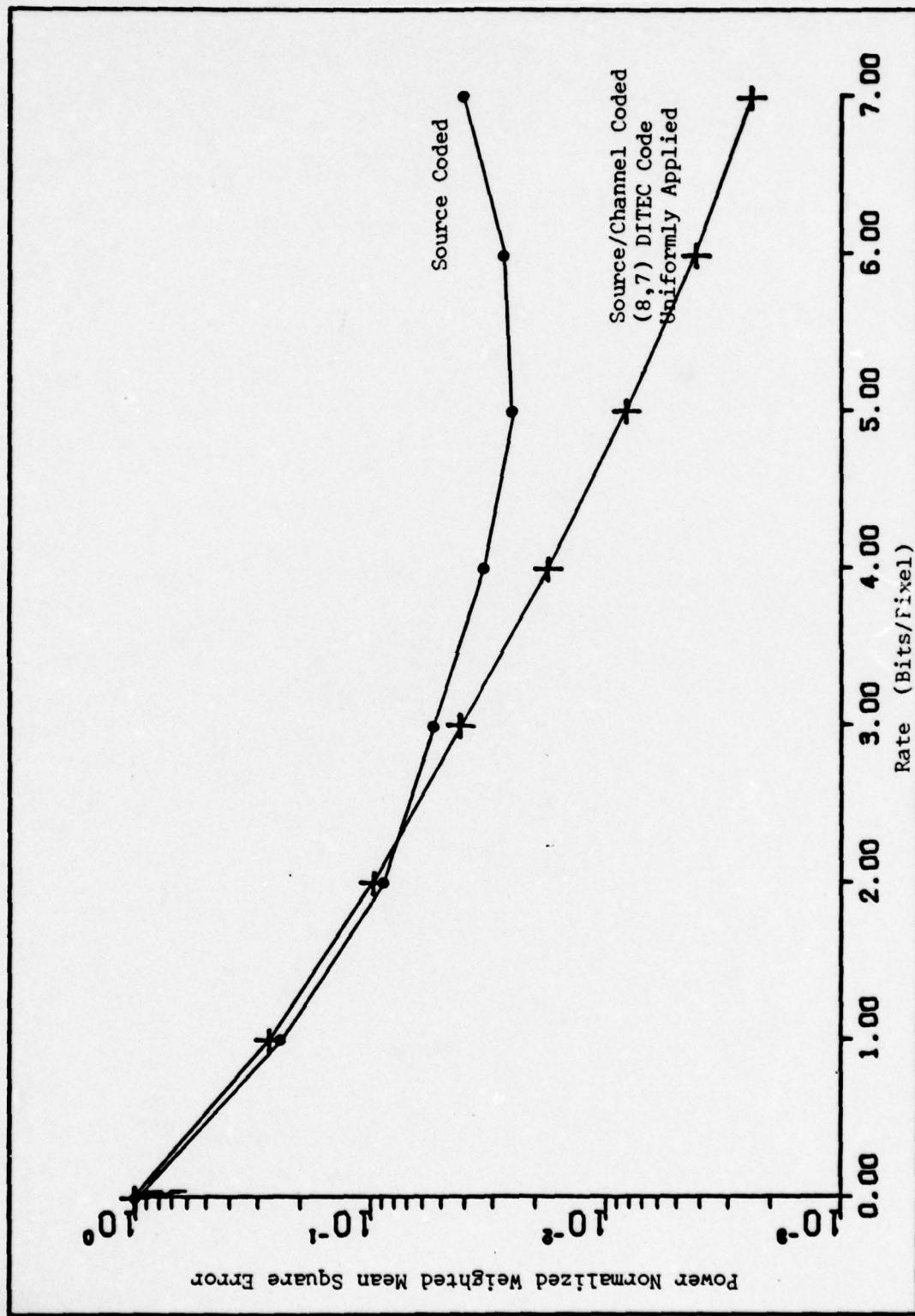


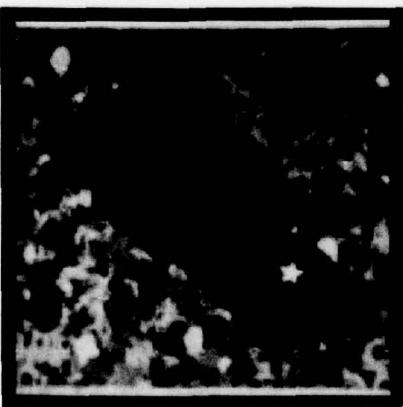
Figure 5.11 Comparison of Source Encoder to Source/Channel Encoder Performance for Simulated Transmissions of Truck Image; $P=10^{-5}$



(a) 6 Bits 3.12% 1.89%
Uncoded



(b) 4 Bits 3.09% 1.81%
Uncoded



(c) 3 Bits 4.10% 2.58%
Uncoded



(d) 6 Bits .41% .28%
(8,7) DITEC Coded



(e) 4 Bits 1.78% 1.36%
(8,7) DITEC Coded



(f) 3 Bits 4.03% 3.10%

Figure 5.12 Comparison of Source_Coded to Source/Channel
Coded Images at P=10
(Distortion Figures below each image are power
normalized WMSE and MSE in that order)

performance at $P=10^{-3}$. The images show large blocks of data to be in error.

This is a result of channel errors in the low frequency DCT coefficients.

Figure 5.11 and 5.12 also show the performance of the simulated image transmission system when the data stream is channel encoded with the (8,7) DITEC code. This was the code that provided the best overall performance in the theoretical calculations. In the simulated transmissions, use of this code provides lower measured distortion than that of the source encoder for all rates down to 3 bits/pixel. The sample images show a clear improvement in quality.

The theoretical results for the (8,7) code predicted that the cross-over in performance between the source coder and the source/channel encoder would occur between 1 and 2 bits/pixel. The relatively poor performance of the simulated source coder causes this crossover point to be pushed out to higher rates. Until the rate is increased to 3 bits/pixel, the quantization error dominates in the total measured distortion. Only above this rate can reductions in the noise error by channel encoding result in better overall performance.

The results of the (8,7) code application were confirmed by all other simulated source/channel encoding combinations. Using the theoretical results as a guide, three additional source/channel combinations were simulated. In the first, the (7,4) Hamming code was applied uniformly to all bits. The second combination protected only a 6×6 square block of the lowest frequency DCT coefficients with the (7,4) Hamming code. The final combination applied the (2,1) Mariner code to the most significant bit of each quantizer output word.

The measured distortion values for all three appear in Figure 5.13.

All three provided better performance than the source encoder alone at rates

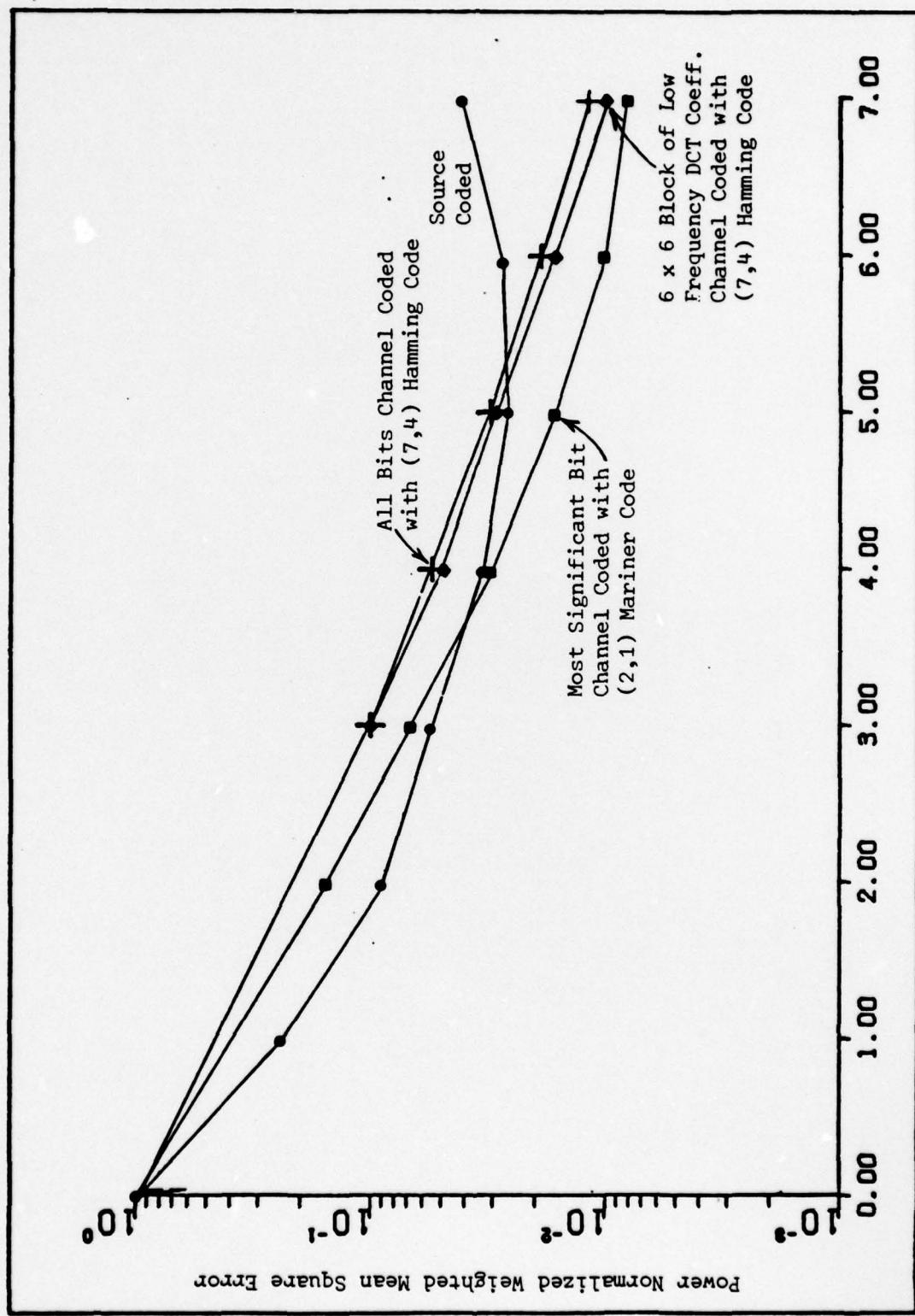


Figure 5.13 Comparison of Source Encoder to Source/Channel Encoder Performance for Simulated Transmissions of Truck Image; $P=10^{-3}$

above 4 bits/pixel. Again, this rate is higher than the theoretically predicted crossover points shown in Figures 4.4, 4.6 and 4.13.

Figure 5.14 presents 2 sample images from each of the three source/channel encoder combinations. Comparing these images to the source encoded images of Figure 5.12, it is difficult to say that the quality has been improved, even though the measured distortion has been decreased. As an example, the (2,1) coded, 6 bit image (Figure 2.14e) has a measured distortion of .93%. The corresponding image for the source encoder (Figure 4.12a) has a distortion of 3.12%. Yet, it is not clear that the quality of the (2,1) coded image is significantly better. Despite these counter examples, the overall simulation results show that the measured distortion does have a strong correlation with perceived image quality. Therefore, comparison of the source/channel encoder performance to the source encoder performance based on either WMSE or MSE is valid.

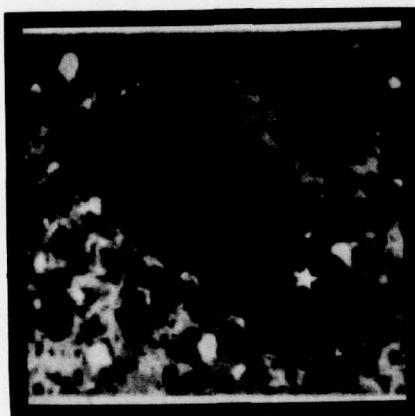


0 Bits 1.89% 1.44%

(7,4) Coded Over Entire Image

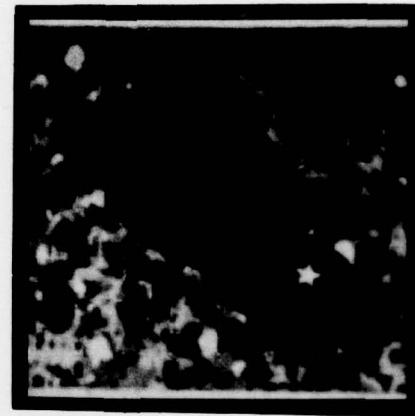


4 Bits 5.52% 4.07%

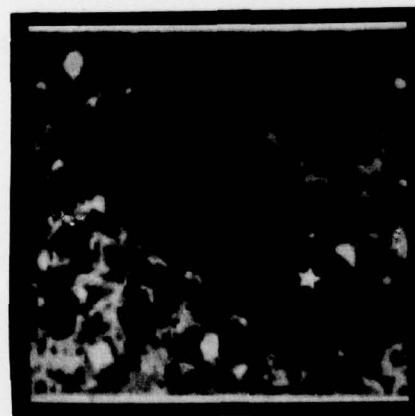


6 Bits 1.82% 1.35%

(7,4) Coded in 6 x 6 Block of DCT Coefficients



4 Bits 5.03% 3.66%



6 Bits .93% .59%

(2,1) Coded on MSE of Each Quantized Coefficient



4 Bits 2.97% 2.48%

Figure 5.14 Source/Channel Encoded Imagery ($P=10^{-3}$)
(Distortion figures under each image are power normalized WMSE and MSE in that order)

VI. Conclusions and Recommendations

Performance Improvement Through the Application of Channel Codes

The primary conclusion which can be drawn from this research is that transmitting imagery, using combinations of source and channel encoding techniques, can result in significant performance improvements. Both the theoretical and simulated results confirm that the measured distortion in images, which have been source and channel encoded prior to transmission, can be lower than the distortion in images which have only been source encoded. The actual magnitude of the performance gain is dependent on a number of factors.

The first of these factors is the transmission rate, that is, the average number of bits allotted to each transmitted pixel. All the theoretical and simulated results showed a "break even" rate between the source encoder's performance and the source/channel encoder's performance. Below the break even rate, the best performance is achieved by transmitting only source bits. Above the break even rate, the best performance is achieved by replacing some source bits with error correction coding bits. In general, as the rate is increased beyond the break even point, the potential gain from the use of channel encoding grows steadily.

The second factor which effects the magnitude of the performance gain is the rate of the channel code (N/K). The theoretical calculations showed that the primary requirement for the channel code is to provide an effective bit error rate of 10^{-5} or less. For error rates in this range, the source encoder provides good performance. Of the codes that can provide error rates in this range, the ones that have a code rate close to 1 provide the

lowest distortion levels in the received image. This is an obvious result, since the low rate codes require very few source bits to be traded off for error correction bits. The effect of this result is to drive the choice of channel codes toward long, low rate block codes or long constraint length, low rate convolutional codes.

A third factor which affects the magnitude of the performance gain is the method used to apply the channel code. If long, low rate codes can be used, the best procedure is to apply them uniformly to all source bits. However, if hardware constraints require short channel codes, better performance may be achieved by only protecting specific subsets of the source bits. For the DCT/block quantization source encoder, improved performance was achieved by channel encoding the low frequency DCT coefficients or the most significant bits of the quantized data.

A final factor which effects the magnitude of the performance gain is the communication channel error rate. The sensitivity of the performance results to this particular factor was not explicitly evaluated. However, some general conclusions can be drawn from the research that was conducted. First, if the error rate of the channel is less than 10^{-5} , there would be no reason to use channel coding. Second, as the error rate increases from 10^{-5} , the benefits of channel encoding grow steadily. Third, if the error rate is greater than 10^{-3} , channel encoding is an absolute necessity. Several source encoded images were transmitted during the simulation at error rates of 5×10^{-3} and 10^{-2} . The received images were totally unrecognizable.

Weighted Mean Square Error Versus Mean Square Error. The research showed little reason to prefer the WMSE distortion criterion to the MSE measure in comparing the relative performance of various source/channel

encoding techniques. Figures 4.14 and 4.15 show that either measure provides essentially the same information. Furthermore, the simulation results indicate that the WMSE criterion may suffer from some of the same ambiguities as the MSE criterion in measuring the quality of imagery. This point would have to be studied at much greater length, however, to make a definite conclusion. Finally, the WMSE criteria greatly complicates the mathematical analysis of imagery transmission systems. Most of the mathematical theory associated with imagery compression and transmission is based on MSE analysis. These results can not always be easily modified to include a weighting function, even one as simple as that used in this report. Definitive proof will have to be offered which shows that the WMSE criterion is significantly better than the MSE criterion in measuring image quality to overcome the analytical benefits of using MSE.

Recommendations for Further Study. The following extension of this research are offered as potential topics for further study.

First, much of the Air Force's interest in imagery transmission is based on the need to transmit video data. Therefore, this research could be extended to include finding procedures for source/channel encoding video signals. Several 3-dimensional source encoding techniques have been proposed in the literature to take advantage of the redundancy between frames as well as within frames of video data.

Second, this research concentrated solely on the discrete cosine transform source encoder. The latest research literature discusses source encoding techniques which achieve greater compression levels than the system used in this thesis. A potential topic for further study would be to examine the sensitivity of the results derived here to changes in the source encoding procedure. Specifically, as the number of source bits transmitted is reduced down to 1/2 bit/pixel and below can the tradeoffs between source

bits and channel coding bits still be made?

The final recommendations are concerned with the AFAL Image Processing System. This facility was found to be invaluable during the course of this research. Unfortunately, since the departure of Dr. Peter Camana (the original sponsor for this thesis) from AFAL, no one in the Lab appears to be using the Image Processing System. In fact, at the time of this writing, this author is currently in possession of most of the operation manuals for the system. Under current Air Force equipment accounting procedures, it will not be long before various pieces of the system (such as the DICOMED or the CONRAC video display) are declared to be surplus. Also, the software for the system is written under a PDP-11 operating system (DOS) which has been superseded, and may not be supported, by AFAL in the future. Therefore, the following recommendations are made to insure the continued availability of the Image Processing System for future AFIT research.

First, as a bare minimum, the system software should be updated to the most current version of the PDP-11 operating system (RSX-11). This work was initiated under contract several years ago, but was never successfully completed. As an indication of the problems associated with using the older operating system, the large PDP-11 disk unit can be used by the Image Processing System only after reformatting the file structure on the disk. The DOS and RSX-11 file formats are not compatible. Other users of the PDP-11 do not appreciate such reformatting of the system disk.

The second recommendation is for AFIT to consider requesting that AFAL donate those portions of the system which are unique to image processing (such as the DICOMED and the CONRAC display) and to establish its own image processing facility. This would be an ambitious project, to say the least.

It appears, however, that the Air Force will have continuing requirements for imagery transmission. Under these circumstances, it would be prudent for AFIT to preserve what little capability is available to perform imagery transmission research.

Bibliography

1. Ahmed, N., T. Natarajan and K.R. Rao. "Discrete Cosine Transform," IEEE Transactions on Computers, C-23: 90-93 (January 1974).
2. Anastassiou, Dimitris and David J. Sakrison. "New Bounds to R(D) for Additive Sources and Applications to Image Encoding," IEEE Transactions on Information Theory, IT-25: 145-155 (March 1979).
3. Andrews, Harry C. and Charles F. Hall. Digital Color Image Compression in a Perceptual Space. AFAL-TR-78-80. Wright-Patterson AFB, Ohio: Air Force Avionics Laboratory, June 1978.
4. Ash, R.A. Information Theory. New York: Wiley-Interscience, 1965.
5. Berger, Toby. Rate Distortion Theory, A Mathematical Basis for Data Compression. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1971.
6. Campbell, F.W. and J.G. Robson. "Application of Fourier Analysis to the Visibility of Gratings," Journal of Physiology, 197: 551-566 (1968).
7. Carl, Joseph W. unpublished notes. School of Engineering Air Force Institute of Technology, Wright-Patterson AFB, Ohio.
8. Chen, Wen-Hsiung and C. Harrison Smith. "Adaptive Coding of Monochrome and Color Images," IEEE Transactions on Communications, COM-25: 1285-1292 (November 1977).
9. Chen, Wen-Hsiung, C. Harrison Smith and S.C. Fralick. "A Fast Computational Algorithm for the Discrete Cosine Transform," IEEE Transactions on Communications, COM-25: 1004-1009 (September 1977).
10. Daut, David G. Rate Distortion Function for 2-D Image Model. Unpublished report. Electrical and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, New York; April 10, 1978.
11. Davisson, Lee D. Rate Distortion Theory and Application. Proceedings of the IEEE, 60: 800-808 (July 1972).
12. Depalma, J.J. and E.M. Lowry. "Sine Wave Response of the Visual System, II. Sine-Wave and Square-Wave Contrast Sensitivity," Journal of the Optical Society of America, 52: 328-335 (March 1962).
13. Farrel, J. and J. Booth. Design Handbook for Imagery Interpretation Equipment. Boeing Aerospace Corporation Report D180-19-063: 1975.
14. Gallagher, Robert. Information Theory and Reliable Communication. New York: John Wiley and Sons, 1968.

15. Gonzales, Rafael C. and Paul A. Wintz. Digital Image Processing. Reading, Massachusetts: Addison-Wesley Publishing Co., 1977.
16. Habibi, Ali. "Two-Dimensional Bayesian Estimates of Images," Proceedings of the IEEE, 60: 878-883 (July 1972).
17. Habibi, Ali and Paul A. Wintz. "Image Coding by Linear Transformation and Block Quantization," IEEE Transactions on Communication Technology COM-19: 50-62 (February 1971).
18. Hall, Charles F. and Ernest L. Hall. "A Nonlinear Model for the Spatial Characteristics of the Human Visual System," IEEE Transactions on Systems, Man and Cybernetics, SMC-7: 161-170 (March 1977).
19. Huang, T.S., Editor. Picture Processing and Digital Filtering. Berlin: Springer-Verlay, 1975.
20. Huang, J.J.Y. and P.M. Schultheiss. "Block Quantization of Correlated Gaussian Random Variables," IEEE Transactions on Communications Systems, CS-11: 289-296 (September 1963).
21. Kurtenback, Aelred J. and Paul A. Wintz. "Quantizing for Noisy Channels," IEEE Transactions on Communication Technology, COM-17: 291-302 (April 1969).
22. MacWilliams, Jessie and Neil Sloane. The Theory of Error Correcting Codes. Amsterdam: North-Holland Publishing Co., 1977.
23. Marnos, James L. and David J. Sakrison. "The Effects of a Visual Fidelity Criterion on the Encoding of Images," IEEE Transactions on Information Theory, IT-20: 525-536 (July 1974).
24. Max, Joel. "Quantizing for Minimum Distortion," IRE Transactions on Information Theory, IT-6: 7-12 (March 1960).
25. McEliece, Robert J. The Theory of Information and Coding. Reading, Massachusetts: Addison-Wesley Publishing Co., 1977.
26. Modestino, J.W. and David G. Daut. Combined Source-Channel Coding of Images. TR 78-1, Electrical and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, New York, September 1978.
27. Natarajan T. Raj and N. Ahmed. "Performance Evaluation for Transforming Coding Using a Non Separable Covariance Model," IEEE Transactions on Communications, COM-26: 310-312 (February 1968).
28. O'Neal, John B. and T. Raj Natarajan. "Coding Isotropic Images," IEEE Transactions on Information Theory, IT-23: 697-707 (November 1977).
29. Papoulis, Athanasios. Probability, Random Variables and Stochastic Processes. New York: McGraw-Hill Book Co., 1965.

30. Pearl, Judea, Harry C. Andrews and William K. Pratt. "Performance Measures for Transform Data Coding," IEEE Transactions on Communications, COM-20: 411-415 (June 1972).
31. Peterson, Wesley and Edward Weldon. Error-Correcting Codes, (Second Edition). Cambridge, Massachusetts: M.I.T. Press, 1972.
32. Pratt, William K. Digital Image Processing. New York: John Wiley and Sons, 1978.
33. Ratliff, Floyd. Mach Bands: Quantitative Studies on Neural Networks in the Retina. San Francisco: Holden-Day, Inc., 1965.
34. Ready, P.J. and Paul A. Wintz. Multispectral Data Compression Through Transform Coding and Block Quantization. Information Note 050572, Laboratory for Applications of Remote Sensing, Purdue University, Lafayette, Indiana, May 1972.
35. Roese, John A., William K. Pratt and Guner S. Robinson. "Interframe Cosine Transform Image Coding," IEEE Transactions on Communications, COM-25: 1329-1339 (November 1977).
36. Sakrison, David J. and V. Ralph Algazi. "Comparison of Line-by-Line and Two Dimensional Encoding of Random Images," IEEE Transactions on Information Theory, IT-17: 386-398 (July 1971).
37. Shannon, Claude E. "A Mathematical Theory of Communication," Bell System Technical Journal, 27: 379-423, 623-656 (1949).
38. Shannon, Claude E. "Coding Theorems for a Discrete Source with a Fidelity Criterion," IRE National Convention Record, Part 4: 142-163 (1959).
39. Stuller, John A. and Bernd Kurz. "Interframe Sequential Picture Coding," IEEE Transactions on Communications, COM-25: 485-495 (May 1977).
40. Tescher, A.G. Transform Image Coding. SAMS-TR-78-127. Los Angeles Air Force Station, Los Angeles, California: Space and Missile Systems Organization, May 1978.
41. Van Ness, F.L. and M.A. Bouman, "The Effects of Wavelength and Luminance on Visual Modulation Transfer," Excerpta Medica International Congress, Series 125: 183-192 (1965).
42. Viterbi, Andrew J. and Jim K. Omura. Principles of Digital Communication and Coding. New York: McGraw-Hill Book Co., 1979.
43. Wintz, Paul A. "Transform Picture Coding," Proceedings of the IEEE, 60: 809-820 (July 1972).

44. Wintz, Paul A. and A.J. Kurtenbach. "Waveform Error Control in PCM Telemetry," IEEE Transactions on Information Theory, IT-14: 650-661 (September 1968).
45. Wu, W.W. "New Convolutional Codes-Part I," IEEE Transactions on Communications, COM-23: 942-952 (September 1975).

Vita

Robert A. Duryea was born on 13 February 1951 in South Egremont, Massachusetts. He graduated from Mount Everett Regional High School in 1968. He attended Rensselaer Polytechnic Institute from which he received a Bachelor of Science in Electrical Engineering in June 1972 and a Master in Business Administration in August 1973. Upon graduation he was commissioned by the USAF through the ROTC program. He received training as a Communications/Electronics Engineer at the Keesler Technical Training Center and was then assigned to the 1844 Engineering and Installation Squadron of Air Force Communications Service at Griffiss AFB, New York. In September 1975 he was assigned to Rome Air Development Center as an Electronic Engineer and continued in the assignment until entering the School of Engineering, Air Force Institute of Technology in June 1978.

Permanent Address: Cooper Hill Road

Ashley Falls, Massachusetts

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 68 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

20. (continued)

Cosine Transform. This technique is combined with both block and convolutional channel codes to form the source/channel encoded imagery transmission system. The distortion in the received image is measured using a weighted mean square error criterion in which the weighting is related to the frequency sensitivity characteristics of the human visual system.

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)